Università degli Studi di Roma "Tor Vergata"

Facoltà di Economia

Corso di laurea in

Economia dei Mercati e degli Intermediari Finanziari

Tesi in Asset Pricing

“Dynamic asset allocation under regime switching: an in-sample and out-of-sample study under the Copula-Opinion Pooling framework”

Il Relatore
Chiar.mo Prof. Stefano Herzel

Il laureando
Andrea Bartolucci

Anno accademico 2015/2016

INDEX

INTRODUCTION

CHAPTER 1 – LITERATURE OVERVIEW

CHAPTER 2 – DATA

2.1 Data sources

2.2 General descriptive statistics

2.3 Time series description

2.4 A comparative analysis of the asset classes returns

2.5 Stationarity and normality tests

CHAPTER 3 – MODEL ESTIMATION

3.1 The basic idea behind regime switching models

3.2 Filtered and smoothed regime probabilities

3.3 The estimation process of a Markov regime switching model

3.4 Overview of the most common multivariate Markov regime switching models

3.4.1 MMSIAH(k,p)

3.4.2 MMSIA(k,0)

3.4.3 MMSIAH(k,0)

3.4.4 MMSIA(k,p)

3.4.5 MMSIH(k,p)

3.4.6 MMSI(k,p)

3.4.7 Multivariate restricted (common underlying Markov chain) MSVAR(K,1) model

3.5 Choice of model specification

3.5.1 Alternative models estimation results

3.5.2 multivariate restricted MSVAR(k,1) models estimation results

3.6 Selected model estimates

3.7 Model restriction test

3.8 Model description

3.8.1 Economic interpretation of regime

3.8.2 Regimes and predictability from the dividend yield

3.8.3 A comparative analysis between the estimated regimes and the NBER USA recession indicator

3.8.4 A dynamic correlation analysis of asset classes returns

3.8.5 Regime shifts in the asset classes returns means and volatilities

3.9 A single regime VAR(1) model estimates

CHAPTER 4 - ASSET ALLOCATION

4.1 Unconditional and state conditional asset classes returns distributions and efficient frontiers based on a MSVAR(2,1) model

4.2 In-sample asset allocation exercise

4.2.1 In-sample asset allocation exercise based on a MSVAR(2,1) model

4.2.2 In-sample asset allocation exercise based on a VAR(1) model

4.3 Out-of-sample asset allocation exercise

4.3.1 Out-of-sample asset allocation exercise based on a MSVAR(2,1) model

4.3.2 Out-of-sample asset allocation exercise based on a VAR(1) model

4.4 Forecasting ability comparison

4.4.1 RMSE

4.4.2 Pearson correlation coefficient

4.4.3 Regression analysis

4.5 Asset allocation exercise conclusions

CHAPTER 5 - ASSET ALLOCATION BASED ON THE COPULA-OPINION POOLING APPROACH

5.1 The copula-opinion pooling (COP) approach

5.2 Out-of-sample mean-variance asset allocation exercise based on COP and MSVAR(2,1) and VAR(1)

5.3 Out-of-sample conditional value-at-risk (CVaR) asset allocation exercise based on COP and MSVAR(2,1) and VAR(1)

5.4 COP asset allocation exercise conclusions

CHAPTER 6 - CONCLUSIONS

INTRODUCTION

Most studies assume that asset returns are generated by a linear process with stable coefficients so the predictive power of variables such as dividend yield does not vary over time. However, there is mounting empirical evidence that asset returns follow a more complicated process with multiple regimes, each of which is associated with a very different distribution of asset returns. Since when authors have started to report evidence of regimes in stock or bond return, regime switching models have been assuming a central role in financial applications because of their well-known ability to capture the presence of rich non-linear patterns in the joint distribution of asset returns. A normal distribution is not sufficient to characterizes asset returns, as historical data showed fat tails, skewness, and kurtosis. Furthermore, a simple model with IID probability distributions, ignoring time series structure of economic stages, seems to be incomplete, as the financial market changes patterns over time. Regime switching models are designed to capture discrete changes in the economic mechanism that generate the data. Usually, regime categorization is linked to the market economic situations. With constant parameters and linear relationships between asset return and factors, the classic asset pricing models do work when the financial market operates normally. However, since authors have displayed that financial crises are characterized by a significant increase in correlations of stock price movements, the benefits of portfolio diversification may seriously be subjected to a reconsideration. As in a standard portfolio optimization model, expected asset returns and their variance and covariance matrix are derived from an asset pricing model. If the model do not takes into account these returns features they are prone to fail to predict expected asset returns correctly since they depend on static parameters. Consequently, the limitations of the classic asset pricing models cause a challenge to portfolio optimization. In this context traditional portfolio optimization models have become unreliable, and the financial industry needs a dynamic portfolio optimization model that is able to characterize different economic conditions. Since the performance of any investment portfolio depends on the accuracy of forecast of assets, the first step is to develop an asset pricing model for the prediction of asset returns which is able to take into account relevant

market returns features, while the second step is to develop a dynamic portfolio model that maximizes the investment profit with a limited risk exposure.

Dynamic asset allocation is a process of selecting instruments and constructing optimal portfolios over time. In this context my work proposes an investment model incorporating market regimes which characterizes different patterns of asset returns in the unobservable economic situations, such as bear and bull markets. The developed model is a dynamic vector autoregressive regime-switching model which also incorporates a predictive variable. The model choose is the result of an extensive model research amongst the Markov Switching Vector Autoregressive model MSVAR which is a general class of models that nests the standard VAR model but additionally accounts for nonlinear regime shifts. The model estimated is employed to model returns of US small stocks, large stock and bonds, and to identify different market regimes over time. From the application of the model to the data, two distinct regimes have been identified. Regime 1 is an extremely persistent bull regime characterized by low realized volatility and positive average realized excess returns on all assets. Regime 2 is a not-highly persistent bear regime characterized by high volatility and large and negative average realized excess returns on small and large stocks while the average realized excess return on bonds are significantly positive and relatively larger than in regime 1. Regime 2 includes two oil shocks in the 1970s, the recession of the early 1980s, the October 1987 stock market crash, the Kuwait invasion in the early 1990s and the ‘Asian flu’ (1996-1998), the dot-com market crash of 2000-2001 (the recent bear market of 2002-2002) and the 2008-2009 financial crisis. In order to evaluate the potential benefit of a multiple regimes model against a more simple single regime model, I additionally estimated a less sophisticated single state model, namely a first order vector autoregressive model VAR(1). The regime switching model covers time variation not only in the conditional means of stock returns and the dividend price ratio, but also in their volatilities, their correlations, and their predictive relation. At each time the returns are assumed to be drawn from one of the different Gaussians distribution underlying the regime switching model. The latent variable is an unobserved state variable that governs the switches between regimes and is assumed to follow a Markov chain. In practice, the Markov chain is unobservable

and embedded in the observed sample over time. Regimes switching model can be seen as mixtures of normal distributions which are considered a very flexible family that can be used to approximate numerous other distributions. Another attractive feature of regime switching models is that they are able to capture nonlinear stylized dynamics of asset returns in a framework based on linear specifications, or conditionally normal distributions, within a regime. This makes asset pricing under regime switching analytically tractable. In particular, regimes introduced into linear asset pricing models can often be solved in closed form because conditional on the underlying regime, normality is recovered. The expected returns and their variance-covariance matrix used in the objective function of the optimization have been generated both by the regimes switching asset pricing model and by the single state asset pricing model.

My work consists of a comparative study of the performances of the multivariate regime switching model against the single regime model in terms of portfolio returns in the context of dynamic asset allocation. The study was conducted through the practical application, both in-sample and out of-sample, of the two models under various portfolio optimization approaches. In the first part of the asset allocation exercise I constructed for any asset pricing model, both in-sample and out-of-sample, two dynamic recursive efficient portfolios that maximize the Sharpe among portfolios on the efficient frontier; in the first in-sample dynamic recursive portfolio the budget constraint is opened up to permit between 0% and 100% in the riskless asset while the second one requires fully-invested portfolios whose weights must sum to 1; in addition short selling, thus negative asset class weights, is not allowed. The other three dynamic recursive portfolios that I constructed have been chosen as those that maximize the investor utility function with three different risk aversion coefficient subject to non negative weights and opened upper budget constraint. The second part of the asset allocation exercise focuses only on the out-of-sample period. Here the Copula-Opinion Pooling approach is applied to implement in the asset pricing model views on the asset returns produced by both the single regime model and the more sophisticated regime switching model. The purpose of this section is to investigate and make a comparison of the behavior of the regime switching model and the single state

model in the COP framework in terms of both expected and realized portfolio returns and Sharpe ratio in the context of mean-variance and conditional value-at-risk (CVaR) portfolio optimization. Therefore, in addition to the five recursive optimal portfolios chosen with the same portfolio selection process as in the first part of the asset allocation problem, here using conditional value-at-risk as the risk exposure constraint, I derived the dynamic optimal weights of other five different portfolios equally distributed, in terms of CVaR, along the time dependent efficient frontier. The exercise has been repeated for different values of the confidence in the views.

From the evaluation of the portfolio performances in the empirical analysis, I expect the superiority of the regime switching model to be noticed. The overperformance can be achieved by the more efficient and desirable risk-reward combinations on the state-dependent frontier that can be obtained only by systematically altering portfolio allocations in response to changes in the investment opportunities as the economy switches back and forth among different states. An investor who ignores regimes sits on the unconditional frontier, thus an investor can do better by holding a higher Sharpe ratio portfolio when the low volatility regime prevails. Conversely, when the bad regime occurs, the investor who ignores regimes holds too high a risky asset weight. She would have been better off shifting into the risk-free asset when the bear regime hits. As a consequence, the presence of two regimes and two frontiers means that the regime switching investment opportunity set dominates the investment opportunity set offered by one frontier.

The plan of the paper is structured as follows. Section 1 describes the literature overview on regime-switching models. Section 2 gives information on the data employed in the paper. Section 3 provides the theoretical background for the Markov regime switching models, estimates a range of regime switching models, select the best one and provides an interpretation of the parameter estimates and the resulting regimes. Section 4 examines the in-sample and out-of-sample performances of asset allocation schemes based on the two different asset pricing models. Section 5 examines the out-of-sample performances of asset allocation

schemes based on the two different asset pricing models under the Copula-Opinion Pooling approach. Section 6 concludes.

CHAPTER 1 – LITERATURE OVERVIEW

Most studies assume that asset returns are generated by a linear process with stable coefficients so the predictive power of state variables such as dividend yield, default and term spreads does not vary over time. However, there is mounting empirical evidence that asset returns follow a more complicated process with multiple “regimes”, each of which is associated with a very different distribution of asset returns. Regime switching models were introduced into economics by Goldfield and Quandt (1973) but their breakthrough in economics came with the seminal paper of Hamilton (1989). Ang and Bekaert (2002a, 2002b), Ang and Chen (2002), Garcia and Perron (1996), Gray (1996), Guidolin and Timmermann (2005a; 2005b; 2006a; 2006b; 2006c), Perez-Quiros and Timmermann (2000), Turner, Startz and Nelson (1989) and Whitelaw (2001) all report evidence of regimes in stock or bond returns. Regime switching models have been assuming a central role in financial applications because of their well-known ability to capture the presence of rich non-linear patterns in the joint distribution of asset returns. A normal distribution is not sufficient to characterizes asset returns, as historical data showed fat tails, skewness, and kurtosis. Furthermore, a simple model with IID probability distributions, ignoring time series structure of economic stages, seems to be incomplete, as the financial market changes patterns over time. There are good economic reasons why the equilibrium distribution of asset returns must be involved with economic sentiment regimes. Regime switching models are designed to capture discrete changes in the economic mechanism that generate the data. Usually, regime categorization is linked to the market economic situations. In the seminal work by Hamilton (1989) the regime switching model allows the data to be drawn from two or more possible distributions, where the transition from one regime to another is driven by the realization of a discrete variable (the regime), which follows a Markov chain process. That is, at each point in time, there is a certain probability that the process will stay in the same regime next period. The transition probability may be constant or they may depend on the other variables. In regime switching model the returns can be described with a hidden Markov model (HMM) of Gaussian mixtures with a number of different regimes. At each time the stock return is

assumed to be drawn from one of the different Gaussian distribution. The latent variable is an unobserved state variable that governs the switches between regimes and is assumed to follow a first order Markov chain, i.e. only the most present history of the chain matters and the switching probabilities are constant. In practice, the Markov chain is unobservable and embedded in the observed sample over time. As pointed out by Marron and Wand (1992), mixtures of normal distribution provide a very flexible family that can be used to approximate numerous other distributions. Mixture of normals can also be viewed as a nonparametric approach if the number of state, k, is allowed to grow with the sample size. They can capture skewness and kurtosis in a way that is easily characterized as a function of the mean, variance and persistence parameters of the underlying starts. They can also accommodate predictability and serial correlation in returns and volatility clustering since they allow the first and second moments to follow a step function driven by shifts in the underlying regime process, c.f. Timmermann (2000). Recent papers have emphasized the importance of adopting flexible models capable of capturing even complicated time-varying forms of heteroskedasticity, fat tails and skews in the underlying distribution of returns, see Manganelli (2004), Patton (2004) and Timmermann (2000). The differences in performance measures between the single-state model and the regime switching model could become larger if higher-order moments are taken into consideration as well. Perez-Quiros and Timmermann (2001) provide an excellent study on higher-order moments under regime switching. Any finite state model is best viewed as an approximation to a more complex and evolving data generating process with non-recurrent states (see, e.g., Pesaran, Pettenuzzo and Timmermann (2006)). Other previous researches indicated that a probability distribution with a structural Markov chain is efficient to describe the dynamics of the economic regimes. Hamilton (1989) successfully applied a two-regime hidden Markov model to the U.S. GDP data and characterized changing pattern of the US economy. Cai (1994), Hamilton and Susmel (1994), and Gray (1996) use variations of the standard Markov regime switching model to describe the time series behavior of U.S. short-term interest rates. Bekaert and Hodrick (1993) document regime shifts in major foreign exchange rates. Schwert (1989)

considered that asset returns may be associated with either high or low volatility which switches over time. Liu, Xu and Zhao (2011) showed that the regime switching model is an effective way for linking sector ETF returns to style and macro factors in changing market regimes over time. Whitelaw (2001) constructed an equilibrium model where growth in consumption follows a regime switching process so investors' intertemporal marginal rate of substitution also follows a regime process. Ang and Bekaert (2002) studied an international asset allocation model with regime shifts and examine portfolio choice for a small number of countries. Guidolin and Timmermann (2006; 2008) provide important economic insights on how investments vary across different market regimes. Recently, Jun Tu (2010) provides a Bayesian framework for making portfolio decisions with regime switching and asset pricing model uncertainty. The joint distribution of future economic indicators not only depends on the current observations but also on the regime switching model parameters. Consequently, asset classes time series follow a multivariate mixture of normal distribution with time-varying mixing parameters over time. As regimes are not observable, the probability distribution of regimes must be dynamically updated with newly observed data and the unconditional joint probability distribution of the returns is a multivariate mixture of normals with mixing parameters equal to the prior distribution of the regimes at time *t.* Regime switching models typically identify bull and bear regimes with very different means, variances and correlations across assets, as noticed by Maheu and McCurdy (2000). As the underlying state probabilities change over time this leads to time-varying expected returns, volatility persistence and changing correlations and predictability in higher order moments such as the skew and kurtosis. The degree of predictability of mean and returns can also vary significantly over time in regime switching models – a feature that seems present in stock returns data as noticed by Bossaerts and Hillion (1999). Under the traditional ARCH and GARCH models of Engle (1982) and Bollerslev (1986), changes in volatility are sometimes found to be too gradual and unable to capture, despite the additions of asymmetries and other tweaks to the original GARCH formulations. Hamilton and Susmel (1994) and Hamilton and Lin (1996) developed regime switching version of ARCH dynamics applied to equity returns that allowed

volatilities to rapidly change to new regimes. A version of a regime switching GARCH model was proposed by Gray (1996). There have been many versions of regime switching models applied to vector of asset returns. Ang and Bekaert (2002a) and Ang and Chen (2002) show that regime switching models provide the best fit out of many alternatives models to capture the tendency of many assets to exhibit higher correlations during down markets than in up markets. Ang and Chen (2002) interestingly find that there is little additional benefit to allowing regime switching GARCH effects compared to the heteroskedasticity already present in a standard regime switching model of normals. It has been known for some time that international equity returns are more highly correlated with each other in bear markets than in normal times (see Erb, Harvey and Viskanta (1994); Campbell, Koedijk and Kofman (2002)). Longin and Solnik (2001) recently formally establish the statistical significance of this asymmetric correlation phenomenon. Whereas standard models of time varying volatility (such as GARCH models) fail to capture this salient feature of international equity return data, recent work by Ang and Bekaert (2002a) shows that asymmetric correlations are well captured by a regime switching model. Stock and bond returns are - to a limited extent – predictable (e.g., Campbell (1987), Fama and French (1988; 1989) and Keim and Stambaugh (1986)), their volatility cluster over time (e.g., Bollerslev, Chou, and Kroner (1992) and Glosten, Jagannathan, and Runkle (1993)) and correlations are not the same in bull and bear markets (e.g., Ang and Chen (2002) and Perez-Quiros and Timmermann (2000)). At shorter horizons stock returns are also far from normally distributed and affected by occasional outliers. Campbell and Ammer (1993) and Fama and French (1989) have showed that variables found to forecast stock returns also predict bond returns. Henkel, Martin, and Nardari (2001) capture the time-varying nature of return predictability in a regime switching context. They use a regime switching vector autoregression (VAR) with several predictors, including dividend yields, and interest rate variables along with stock returns. They find that predictability is very weak during business cycle expansions but is very strong during recessions. Thus, most predictability occurs during market downturns, and the regime switching model captures this countercyclical predictability by exhibiting significant predictability only in the contraction regime. A branch of the existing

regime switching literature concern with the issue of parameter estimation. For an extensive overview concerning the econometrics issues of regime switching model and an overview about empirical evidence many authors refer to Kim and Nelson (1999). One of the first papers in financial econometrics that estimates time-varying integration of single countries to the world market is Bekaert and Harvey (1995). Hamilton (1994) and Kim and Nelson (1999) give an overview about the econometrics of state-space models with regime switching and provide an overview of possible applications to finance. From an econometric point of view, the main problem in estimating regime switching models is the unobservability of the prevailing regime. Two different approaches have been suggested: a classical maximum likelihood based on the filters such as the Hamilton filter or on the expectation maximization algorithm and a Bayesian approach based on numerical Bayesian methods such as the Gibbs sample and Markov Chain Monte Carlo methods. Regime probabilities play a critical role in the estimation for regime switching models, which uses maximum likelihood techniques (see Hamilton (1994), Gray (1996); and Ang and Bekaert (2002b)) or Bayesian techniques (see Albert and Chib (1993)). It is well known that applications of classical mean-variance frontier (MVF) technology to dynamic asset allocation problems in which the MVF is allowed to depend on one more variables capturing the state of market investment opportunities, suffer from a number of issues (e.g., see Schöttle an Werner (2006)). For instance, the shape of the MVF together with the location of the efficient portfolios has been observed to change drastically as market data are progressively updated and expanded. Moreover, it is typical to observe that MVFs often occupy rather unrealistic regions of the mean-standard deviation space as a result of optimization based on error-prone estimations, resulting in large deviations between the ex-ante, in-sample and the ex-post, out-of-sample Sharpe ratios. Guidolin and Ria (2010) can be seen as an attempt to produce more robust estimates of the MVF and hence – after appropriate mean-variance preferences have been assumed – more robust optimal portfolios not by changing methods of estimation or by resampling the data, but instead by exploring the implication of a simple and yet powerful parametric approach that explicitly tracks the time variation in the features of the investment opportunity sets (means, variances, and

correlations) as depending from a latent Markov state variable. In the presence of regimes, asset returns may have entirely different relationships with the predictors in different regimes. One of the key improvements upon a traditional investment model is that portfolio decisions are based on a Bayesian type of dynamic updating on the probability distribution of the economic regimes. Essentially, this modeling approach provides time-varying risk premiums and risk magnitudes, depending not only on the economic indicators but also the updated regime distribution at any point in time. Investment securities may exhibit different risk levels in different economic situations and therefore, different risk premiums. However, there is no clear determination as to which economic regimes we are in by directly observing the market data. The key idea for a regime switching model is to resolve the issue of unobserved economic regimes over time. Regime switching means that all conditional moments of the asset returns distribution are time-varying, so it is possible to extend the previous literature on strategic asset allocation to cover the case where all moments may be appropriate. However, none of the state can be perfectly anticipated: starting from any one of these the investor always assigns a positive and non-negligible probability to the possibility of transitioning to a different state. Regime switching models also nest as a special case jump models, given that a jump is a regime that is immediately exited next period and, when the number of regimes is large, the dynamics of a regime switching model approximates the behavior of time-varying parameter models where the continuous state space of the parameter is appropriately discretized. Finally, another attractive feature of regime switching models is that they are able to capture nonlinear stylized dynamics of asset returns in a framework based on linear specifications, or conditionally normal log-normal distributions, within a regime. This makes asset pricing under regime switching analytically tractable. In particular, regimes introduced into linear asset pricing models can often be solved in closed form because conditional on the underlying regime, normality (or log-normality) is recovered. This makes incorporating regime dynamics in affine models straightforward. Regime shifts continue to have a significant effect on the optimal asset allocation and expected utility even after accounting for parameter uncertainty. Guidolin and Timmermann Size and Value Anomalies under Regime

Shifts (2005) find that four-state models perform better than single-state alternatives both in terms of precision of their out-of-sample forecasts and in terms of sample estimates of mean returns and that accounting for the presence of regimes lead to higher average realized utility even after accounting for parameter estimation error. Another branch of literature that concerns with the portfolio choice and regime switching analyzes the effects of regime switching on asset allocation. Overall, the main findings are that regime switching induces a change in the asset allocation depending on the investment horizon and depending on the current regime. One of the main references for asset allocation in regime switching framework is Ang and Bekaert (2002). In their paper, they analyze dynamic asset allocation with regimes shifts in an international context. The starting point of their paper are time-varying correlation between different equity markets. In bad times correlation and volatilities increase in comparison to good times and, therefore the investment opportunity set is stochastic. In the empirical part, they assume two state model with Markov switching and constant transition probabilities. For parameter estimation, they use a Bayesian procedure similar to Hamilton (1989) and Gray (1996). Overall, there are always relatively large benefits of international diversification, although the optimality of the home-biased portfolios cannot always be rejected statistically. The cost of ignoring regime switching are very high if the investor is allowed to switch to cash position. If the investment universe is limited to equities, costs of ignorance are lower. With respect to hedging demands, they find that intertemporal hedging demands under regime switching are economically negligible and statistically insignificant. Similar, Ang and Bekaert (2004) find that for a global all-equity portfolio, the regime switching strategy dominates static strategies in an out-of-sample test. In a persistent high-volatility market, the model tells the investors to switch primarily to cash. Recent contributions include Graflund and Nilsson (2003), Bauer, Haerden and Molenaar (2003), Ang and Bekaert (2004), and Guidolin and Timmermann (2005). A number of authors analyze the implications of regime switching in portfolio selection. The presence of asymmetric correlations in equity returns has so far primarily raised a debate on whether they cast doubt on the benefits of international diversification, in that these benefits are not forthcoming when you need them the most. However, the

presence of regimes should be exploitable in an active asset allocation program. The optimal equity portfolio in the high volatility regimes is likely to be very different (for example more home biased) than the optimal portfolio in the normal regime. When bonds and T-bills are considered, optimally exploiting regime switching may lead to portfolio shifts into bonds or cash when a bear market regime is expected. In particular, investor often use available realized returns at a given point in time to determine whether the market is in a bull or a beat state. Turner, Starts, and Nelson (1989) provide a rigorous econometric model for analyzing bull and bear markets and find that the S&P500 index displays different means and variances across these markets. Schwert (1989) and Hamilton and Susmel (1994) also document regime-dependent market volatility, while Ang and Bekaert (2002; 2004) and Guidolin and Timmermann (2005; 2007; 2008a; 2008b) provide important economic insights on how investments vary across different market regimes. Jun Tu (2010) found that the certainty-equivalent loses associated with ignoring regime switching are generally above 2% per year, and can be as high as 10%. Tu and Zhou (2004) find that the certainty-equivalent losses associated with ignoring fat tails are typically less than 1% for mean-variance investors facing model and parameter uncertainty. However, the findings of this study reveal that ignoring regime switching can lead to sizable economic costs. These findings support the qualitative conclusions of the earlier regime studies by Ang and Bekaert (2002; 2004) and Guidolin and Timmermann (2005; 2007; 2008a; 2008b), despite their classical framework, which does not incorporate model or parameter uncertainty. This is because the impact of these types of uncertainty could be less important than the impact of regimes. Where the investor is allowed to rebalance his or her portfolios during the investment period, the corresponding problem is usually discussed in terms of intertemporal hedging, first mentioned by Merton (1971). The intertemporal hedging is a desire of the investor to protect him from unfavorable changes in the set of investment opportunities, or a desire to be able to profit from favorable changes. Ang and Bekaert (1999) investigates international diversification and intertemporal hedging demands within a regime switching framework from the perspective of a US investor who is allowed to buy foreign stocks in addition to US stocks. Nilsson and

Graflund (2001) instead assume that the investment opportunity set faced by the investor is spanned by a well-diversified home stock market portfolio and risk-free short-term bill. The aim of the authors and contribution of their paper is to study the importance of regime switching in this classical portfolio selection setup. The authors investigated if the optimal portfolio differs across different regimes, and if the intertemporal hedging demand differs across regimes. In the regime switching model described by the authors the returns are represented as a mixture of Gaussian distributions. The joint effects of learning about the underlying state probabilities and predictability of asset returns from the dividend yield give rise to a non-parametric relationship between the investment horizon and the demand for stocks. Strategic asset allocation decisions can only be made in the context of a model for the joint distribution of asset returns. Most studies assume that asset returns are generated by a liner process with stable coefficients so the predictive power of state variable such as dividend yields, default and term spreads does not vary over time. However, there is mounting empirical evidence that asset returns follow a more complicated process with multiple "regimes", each of which is associated with a very different distribution of asset returns. In Guidolin and Timmermann the authors characterize investors' strategic asset allocation and consumption decisions under a regime switching model for asset returns with four states characterized as crash, slow growth, bull and recovery states. A difference to earlier studies is that the authors allow the underlying states to be unobservable to the investor who must infer the state probability from the sequence of returns data. Kao and Shumaker (1999) analyze the opportunities for equity style timing, based on Fama and French (1993) factors, using recursive partitioning (regression and classification trees) and macroeconomic factors (term spread, real bond yield, corporate credit spread, high-yield spread, estimated GDP growth, earnings-yield gap, CPI), they try to predict future differences in style returns. They find that timing strategies in the US market based on asset class and size have historically provided more opportunity for outperformance than a timing strategy based on value and growth. For example, during the Internet bubble period 1998-2002, the stock market was extremely volatile, while market volatility was relatively low in the period of 2003-2006. It is highly probable that market sentiment, market volatility,

and the non-smooth asset return processes are regime dependent. Ignoring such a possibility and simply averaging information across the market regimes may result in suboptimal investment strategies. If regimes exist and may be identified , estimated, and predicted, then it is an open question whether an investor should take notice of them, and go through the relatively sophisticated econometric techniques required by her acknowledging this state-dependence. Clarke and de Silva (1998) note that no static mix to be applied to standard mean-variance portfolios can be used to achieve a point along a state-dependent efficient frontier. The more efficient and desirable risk-reward combinations on the state-dependent frontier may be achieved only by systematically altering portfolio allocations in response to changes in the investment opportunities as the economy switches back forth among different states. The reason is that in the presence of state-dependence (when two states with probabilities p and $1-p$ are possible), a mixture of Gaussian (more generally, elliptical) densities is never the same as a Gaussian density that has means and variances which are probably-weighted (with weights p and $1-p$) averages of the state-dependent means and variances. When investment opportunities remain constant over time, a power utility investor's horizon does not affect the optimal asset allocation, c.f. Samuelson (1969). In the absence of predictor variables, standard models therefore imply constant portfolio weights. In contrast, using the dividend yield as a predictor, Barberis (2000) finds that the weight on stocks should increase as a function of the investor's horizon. Even in the absence of predictor variables, regime switching models imply that investors' asset allocation varies over time as the underlying states offer different investment opportunities and investors revise their beliefs about the state probabilities. Ang and Chen (2002) find that equity correlations that differ across high/low return states can be successfully captured by a regime switching model. They note that small firms' returns exhibit relatively strong asymmetries and argue that such asymmetric correlations may be important for strategic asset allocation purposes, although they stop short of analyzing this question. Guidolin and Timmermann (2005) extend the regimes switching model for asset returns to include predictability from state variables such as dividend yield. Consistent with earlier findings in the literature (e.g., Campbell, Chan and Viceira (2003)), Guidolin

and Timmermann (2005) find that the recursively updated portfolio weights vary significantly over time as a result of changing investment opportunities and that optimal asset holdings are sensitive to how predictability is modeled. When regimes are taken into account, there is evidence that the allocation to stocks and bonds as well as the division of stock holdings among small and large firms is quite different from that obtained under linear models of predictability in asset returns. Furthermore, the authors generally find that the average realized utility is highest for model that account for regime switching. Huy Thanh Vo and Maurer (2013) solve the asset allocation problem under stock return predictability based on the dividend yield for an investor who accounts for both; the uncertainty about the true underlying predictive power of the dividend yield and changes in the joint distribution of stocks and predictors due to shifts in regimes. The model proposed by the authors covers time variation not only in the conditional means of regime switching and the dividend yield, but also in their volatilities, their correlations, and their predictive relation. The possibility of switching across regimes, even if it occurs relatively rarely, induces an important additional source of uncertainty that investors want to hedge against. It is reasonable to expect that if the market portfolio exhibits regime switches, then portfolios of stocks would also switch regimes, and the regimes and behavior within each regimes of the portfolios should be related across portfolios. This is indeed the case Perez-Quiros and Timmermann (2000), Gu (2005), and Guidolin and Timmermann (2008b), among others, fit regime switching models to small cross section of stock portfolios. These studies show that the magnitude of size and value premiums, among other things, varies across regimes in the same direction. On the other hand, the dynamics of certain stock portfolios react differently across regimes, such as small firms displaying the greatest differences in sensitivities to credit risk across recessions and expansions compared to large firms. Factor loadings of value and growth firms also differ significantly across regimes. They exploit the ability of the regime switching model to capture higher correlation during market downturns and examine the question of whether such higher correlations during bear markets negate the benefits of international diversification. They find there are still large benefits of international diversifications. The cost of ignoring regimes is very large

when a risk-free asset can be held; investors need to be compensated approximately two to three cents per dollar of initial wealth to not take into account regimes changes. In the regimes switching context the risk-return trade-off can vary across states in a way that may have strong asset allocation implications. For example, knowing that the current state is a persistent bull state will make most risky assets more attractive than in a bear state. Asset allocation under an unknown number of permanent structural breaks has been studied by Pettenuzzo and Timmermann (2011), who apply a multiple change point model proposed by Chib (1998). When dividend yield as a predictor is added to the a regime switching model the resulting regime switching VAR model nests many of the models in the existing literature and enables the correlation between the dividend yield and asset returns to vary across different regimes. The relationship between stock returns and the dividend yield is linear within a given regimes. However, since the coefficient on the dividend yield varies across regimes, as the regime probabilities change the model is capable of tracking a non-liner relationship between asset returns and the yield. This is important given the evidence of a non-linear relationship between the yield and stock returns uncovered by Ang and Bekaert (2004). A large literature in finance has reported evidence that variables related to the business cycle predicted stock and bond returns. One of the key instruments is the dividend yield; see, e.g., Campbell and Shiller (1988), Fama and French (1998; 1989), Ferson and Harvey (1991), Goetzmann and Jorion (1993) and Kandel and Stambaugh (1996). Due to its high persistence coupled with the strong negative correlation between shocks to returns and shocks to the dividend yield, Campbell, Chan, and Viceira (2003) find that the dividend yield generate the largest hedging demand among a wider set of predictable variables. Empirical studies on predictive regressions find economic indicators related to the business cycle slightly effective in predicting stock returns, which contradicts the assumption of independently and identically distributed stock returns (Campbell, Low, and MacKinlay (1997)). Against the background of these empirical findings, portfolio theory reveals that optimal policies under predictability do exhibit horizon effects, as time variation in the expected returns induce intertemporal hedging demands against changes in the investment opportunity set (Brennan, Schwartz, and Lynch

(1999)). The implications were first formulated by Merton (1973). The lasting stock bull market during the 80s and 90s reinforced doubts about predictability, as the dividend yield ability to forecast expected stock returns declined noticeably. Acknowledging this ongoing dispute, portfolio theory has accounted for uncertainty in predictive relations by incorporating estimation risk (Kandel and Stambaugh (1996), Barberis (2000)), model risk (Avramov (2002)), and learning (Xia (2001), Brand and Santa-Clara (2006)). These studies adopt a Bayesian framework instead of taking a binary view, either accepting or rejecting predictability solely based on statistical significance. Nevertheless, a comprehensive re-examination of predictive regressions using an extended data set and with newly formulated test statistics has led to some researchers to conclude that stock return predictability, especially long run predictability, was a statistical fluke and never truly existed (Ang and Bekaert (2007), Boudoukh, Richardson, and Whitelaw (2008)). Moreover the out-of-sample performance of several predictors has been reported to be poor and, in many cases, even worse, than unconditional mean of stock returns (Bossaerts and Hillion (1999), Welch and Goyal (2008)). This too casts doubt on the practical relevance of predictability. In contrast to these studies, another strand of literature argues that the predictive relation itself is subject to time variations and hence cannot be sufficiently described by a simple linear regression. For instance, Pesaran and Timmermann (1995), Paye and Timmermann (2006), and Henkel, Martin, and Nardari (2011) provide evidence for changes in the predictive relation over time. Lettau and Van Nieuwerburgh (2008) argue that the mean of dividend yield was affected at least by one or even two structural breaks (in 1994 and possibly in 1951), and that once adjusted for the subsample means, the predictive power remains stable and significant. Actually standard regime switching models cannot capture the effects of permanent structural breaks. Nevertheless, they can approximate permanent breaks to a certain extent by including extremely persistent regimes. In the long run, however, they imply a steady state distribution. Furthermore, the combination of regime shifts and stock predictability also incorporates two aspects of major importance in asset allocations. Generally, regime shifts generate a momentum effect, which increases the variance in the long run, whereas stock return predictability generates a mean

reversion, or rebound effect, which decreases the variance in the long run (Ang and Bekaert (2002)). Samuelson (1991) define a “rebound” process, or mean-reverting process, as having a transition matrix which has a higher probability of transitioning to the alternative state than staying in the current state. Samuelson shows that with a rebound process, risk-averse investors increase their exposure to the risky assets as the horizon increases. That is, under rebound, long-horizon investors are more tolerant of risky assets than short-horizon investors. The opposite of rebound process is called “momentum” process: it is more likely to continue in the same state rather than transition to the other state. Under a momentum process, risk-averse investors want to decrease their exposure to risky assets as horizon increases. Intuitively, long-run volatility is smaller under a rebound process that under a momentum process (with the same short-run volatility).

CHAPTER 2 – DATA

In this chapter I am going to describe in more detail the data I have used in this work. In the first paragraph the data are listed and the sources are provided; in the second paragraph some general descriptive statistics are provided for the risk free rate, the dividend yield and the three asset classes time series; in the third paragraph a comment of the data presented in the second paragraph and a general overview are provided for each time series; in the fourth paragraph, the three asset classes time series are compared and some further comments are provided; lastly, in the fifth paragraph the results of the normality and stationarity tests are provided and described for the three asset classes time series and the dividend yield.

2.1 Data sources

Table 1 Data Short Name and Description

Variable	Definition
lo20	excess value weighted monthly returns of first and second CRSP size decile US equity portfolios
hi20	excess value weighted monthly returns of ninth and tenth CRSP size decile US portfolios
tbond	excess monthly returns on a 10 years bond priced using the monthly FED 10-Year Treasury Constant Maturity Rate
risk_free	1-month US T-Bill monthly returns
div_y	S&P 500 Dividend Yield

The time series lo20 and hi20 represent respectively the excess value weighted monthly returns of the first and second, and the ninth and tenth size deciles US equity portfolio. The data are made available by Kenneth R. French on his academic website's data library[1]. According to the authors the portfolios are constructed at the end of each June using the June market equity and NYSE breakpoints. The portfolios for July of year t to June of t+1 include all NYSE, AMEX, and NASDAQ stocks for which market equity data for June of year t were available.

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html and http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_port_form_sz.html

The version of the data I have used dates back to March 2015 and the data author declares the file was created using the January 2015 CRSP (The Center for Research in Security Prices) database.

The time series tbond represents the excess monthly returns on a US 10-year constant maturity Treasury Bond. The time series has been realized pricing a 10-year constant maturity bond using the monthly US 10-year Treasury Constant Maturity Rate provided by the US Federal Reserve Bank of St. Louis. The pricing formula applied is the monthly version of the Damodaran formula[2] which models the bond returns as the monthly fraction of the annual promised coupon at the start of the year and the bond price change due to interest rate changes.

The risk_free time series represents the US 1-month Treasury Bill rate. The data are published by Kenneth R. French on his academic website's data library[3] and are made available by Ibbotson and Associates Inc. The div_y time series represents the S&P 500 Dividend Yield which consists in the annual dividends paid by the companies in the index divided by the price index. The data are made available by Robert Shiller on his academic website[4] in the stock market data section. The same data have been used by the author in his book *Irrational Exuberance* (2005).

2.2 General descriptive statistics

In this paragraph some general descriptive statistics of the four time series are shown below in Table 2.

Table 2 Monthly Time Series General Descriptive Statistics

	lo20	**hi20**	**tbond**	**risk_free**	**div_y**
# of observations	540	540	540	540	540
Mean	0.0068	0.0036	0.0016	0.0046	0.0311
Median	0.0103	0.0065	0.0018	0.0043	0.0308

[2] http://quant.stackexchange.com/questions/3941/t-note-returns-from-t-note-yields-derivation-of-damodarans-formula and http://www.stern.nyu.edu/~adamodar/pc/datasets/histretSP.xls
[3] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html and http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_port_form_sz.html
[4] http://www.econ.yale.edu/~shiller/data.htm

Minimum	-0.3022	-0.2091	-0.0894	0.0000	0.0111
Maximum	0.2732	0.1761	0.1005	0.0135	0.0624
Standard Dev.	0.0650	0.0437	0.0202	0.0023	0.0120
Variance	0.0042	0.0019	0.0004	0.0000	0.0001
Skewness	-0.2177	-0.3726	0.4497	0.8436	0.2917
Kurtosis	5.2645	4.7089	6.1871	4.5912	2.3683
Sharpe Ratio	0.1051	0.0824	0.0776		
Percentile (10%)	-0.0697	-0.0483	-0.0215	0.0015	0.0158
Percentile (25%)	-0.0295	-0.0207	-0.0111	0.0032	0.0203
Percentile (50%)	0.0103	0.0065	0.0018	0.0043	0.0308
Percentile (75%)	0.0442	0.0318	0.0130	0.0057	0.0386
Percentile (90%)	0.0794	0.0536	0.0236	0.0075	0.0489
Mean (normal fit)	0.0068	0.0036	0.0016	0.0046	0.0311
Mean annualized (normal fit)	0.0821	0.0433	0.0188	0.0547	
Standard Dev. (normal fit)	0.0650	0.0437	0.0202	0.0023	0.0120
Standard Dev. Annualized (normal fit)	0.2253	0.1515	0.0699	0.0081	

2.3 Time series description

The US small cap excess returns, hereafter called lo20, has yielded on average a monthly mean return of 0.68%, a standard deviation of 6.5%, which is equivalent to a variance of 0.42%, and therefore a Sharpe ratio of 0.1051; the median is equal to 1.03% and the maximum and minimum sample values are respectively 27.32% and -30.22%. In addition a normal distribution has been estimated from the sample data, the estimation returns an estimate of the mean equal to 0.68%, which is equivalent to an annualized mean of 8.21%, and an estimate of the standard deviation of 6.50%, which is equivalent to an annualized standard deviation of 22.53%. The time series exhibits a negative skewness and kurtosis in excess of the Gaussian benchmark (equal to three) respectively -0.2177 and 5.2645. The lo20 time series is depicted in the plots below.

Figure 1 lo20 time series plot

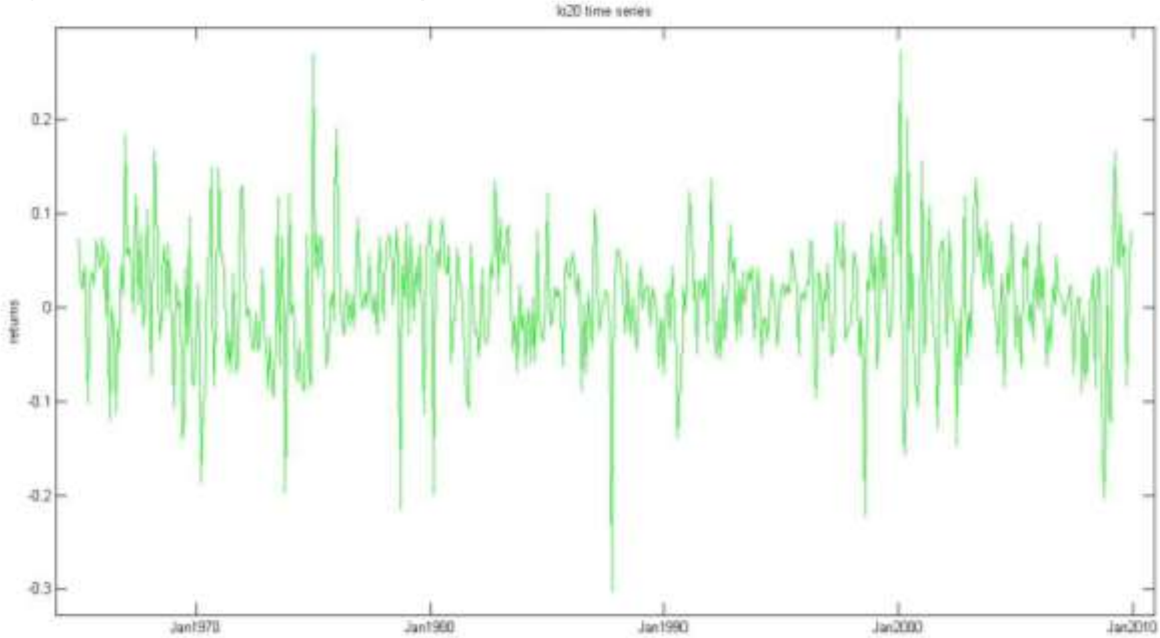
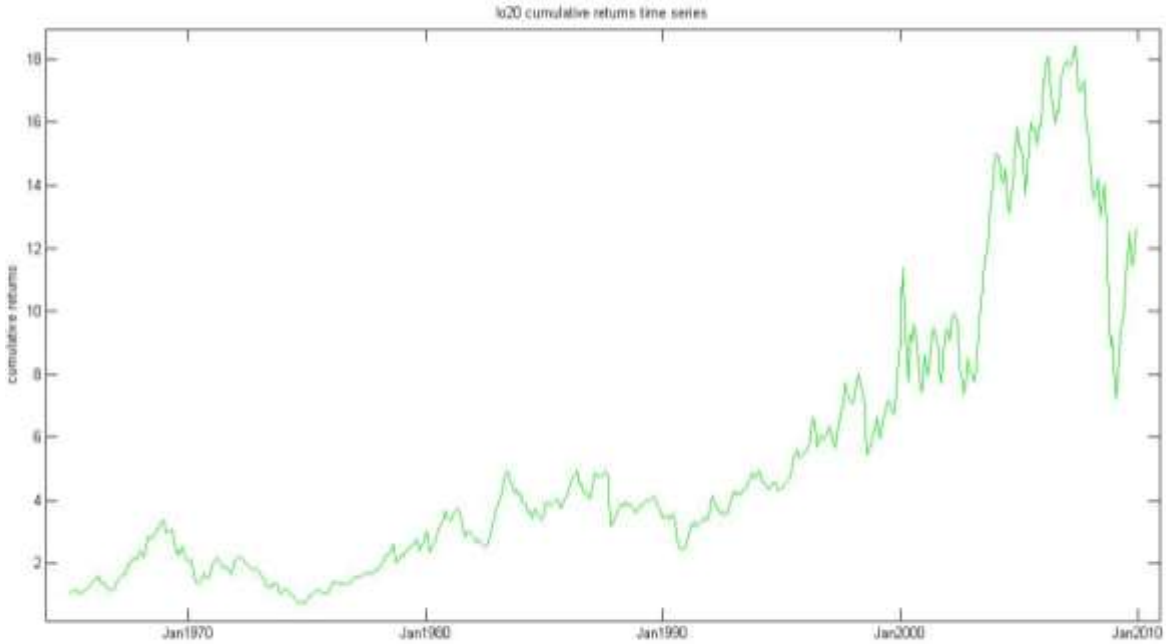


Figure 2 lo20 cumulative returns time series plot



From the two plots above it can be seen that the US small cap excess returns has experienced, over the sample period, several bull and bear phases. From a first glance it can be clearly seen that there is an alternation of high and low volatility periods. Overally the time series is characterized by an uptrend which becomes steeper starting from the beginning of the 1990s.The asset class has overally

performed very well reaching a cumulative return equal to about 11 (1100%) over the sample period. The peak in the evolution of a unit of wealth invested in lo20 has been reached around 2007, just before the last financial crisis, at a level of about 18 (1700% cumulated return).

The US large cap excess returns, hereafter called hi20, has yielded on average a monthly mean return of 0.36%, a standard deviation of 4.37%, which is equivalent to a variance of 0.19%, and therefore a Sharpe ratio of 0.0824; the median is equal to 0.65% and the maximum and minimum sample values are respectively 17.61% and -20.91%. In addition a normal distribution has been estimated from the sample data, the estimation returns an estimate of the mean equal to 0.36%, which is equivalent to an annualized mean of 4.33%, and an estimate of the standard deviation of 4.37%, which is equivalent to an annualized standard deviation of 15.15%. The time series exhibits a negative skewness and kurtosis in excess of the Gaussian benchmark (equal to three) respectively -0.3726 and 4.7089. The hi20 time series is depicted in the plot below.

Figure 3 hi20 time series plot

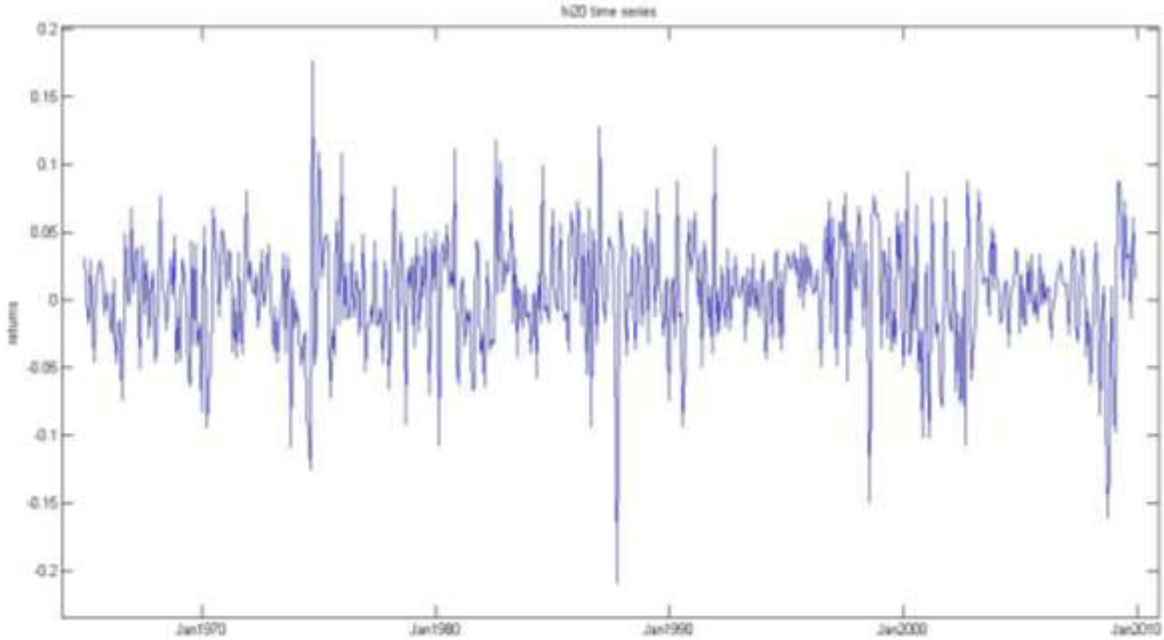
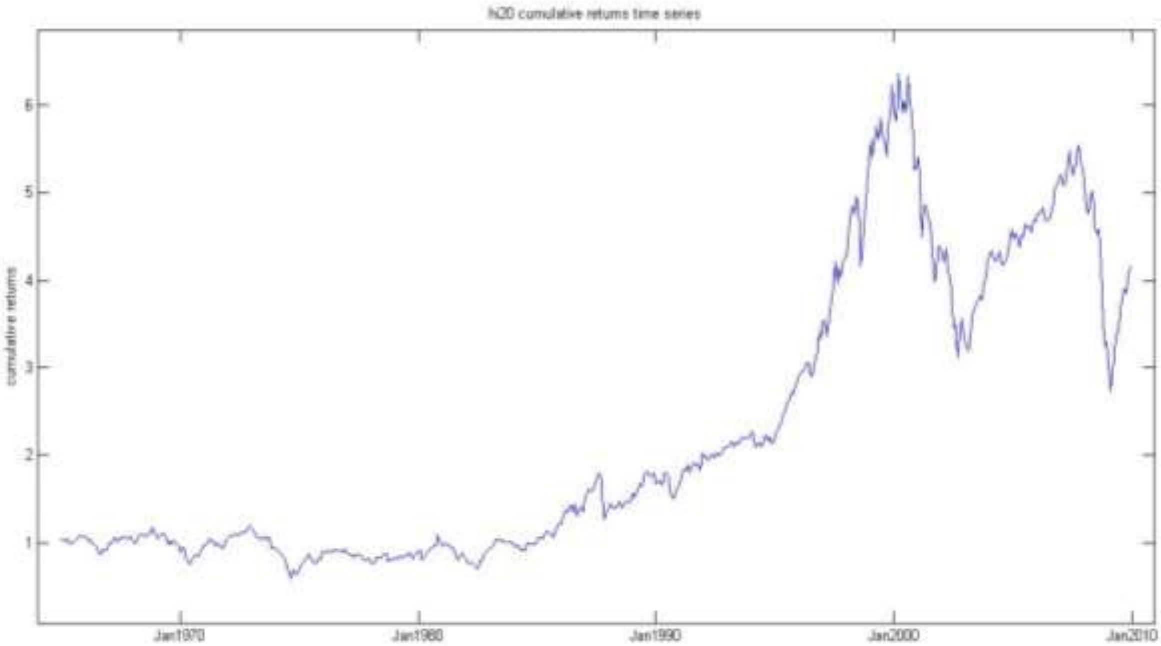


Figure 4 hi20 cumulative returns time series plot



From the two plots above it can be seen that the US large cap excess returns has experienced, over the sample period, several bull and bear phases, similar to what happened to lo20. From a first glance it can be clearly seen that there is a succession of high and low volatility periods. Overally the time series is characterized by an uptrend which becomes steeper starting from the beginning of the 1990s and much steeper from the second half of the 2000s.The asset class has overall performed slightly well reaching a cumulative return equal to about 3 (300%) over the sample period. The peak in the evolution of a unit of wealth invested in hi20 has been reached around January 2000, just before the dot-com burst, at a level slightly above 6 (500% cumulated return).

The 10-Year US Treasury Bond excess returns, hereafter called tbond, has yielded on average a monthly mean return of 0.16%, a standard deviation of 2.02%, which is equivalent to a variance of 0.04%, and therefore a Sharpe ratio of 0.0776; the median is equal to 0.28% and the maximum and minimum sample values are respectively 10.05% and -8.94%. In addition a normal distribution has been estimated from the sample data, the estimation returns an estimate of the mean equal to 0.16%, which is equivalent to an annualized mean of 1.88%, and an estimate of the standard deviation of 2.02%, which is equivalent to an

annualized standard deviation of 6.99%. The time series exhibits a positive skewness and kurtosis in excess of the Gaussian benchmark (equal to three) respectively 0.4497 and 6.1871. The tbond time series is depicted in the plot below.

Figure 5 tbond time series plot

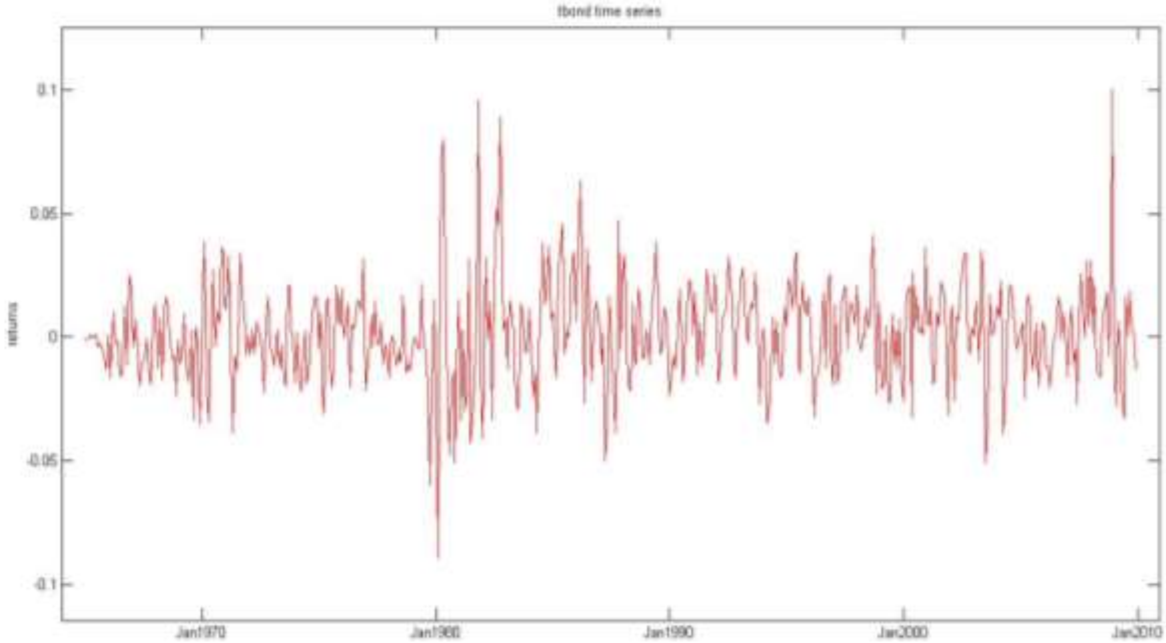
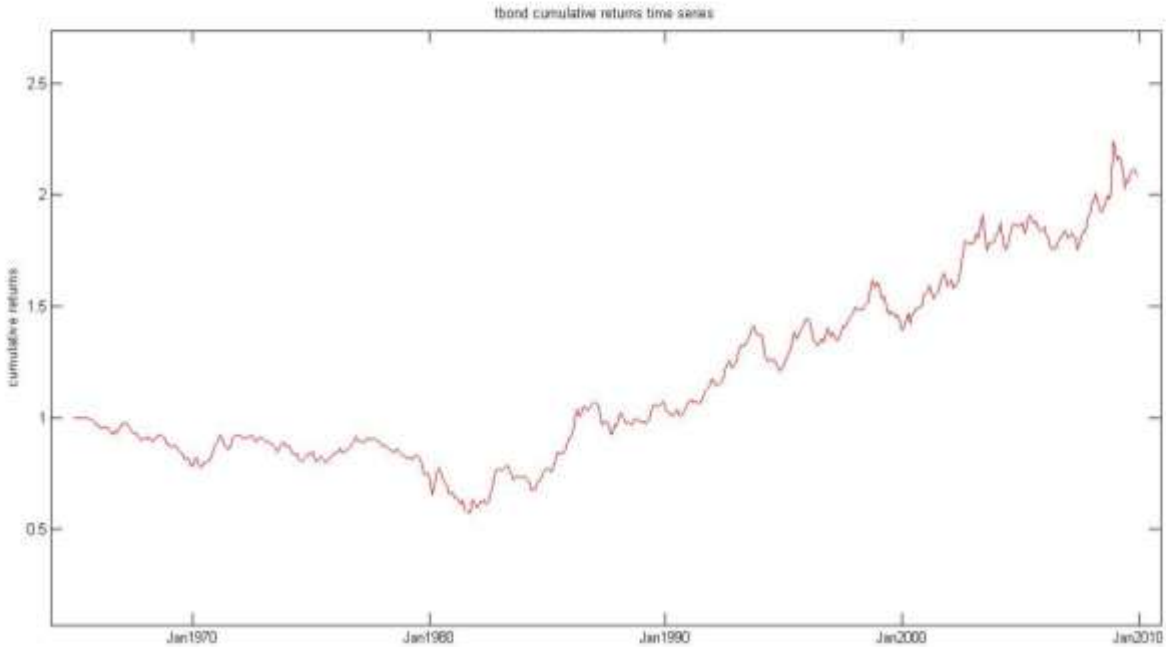


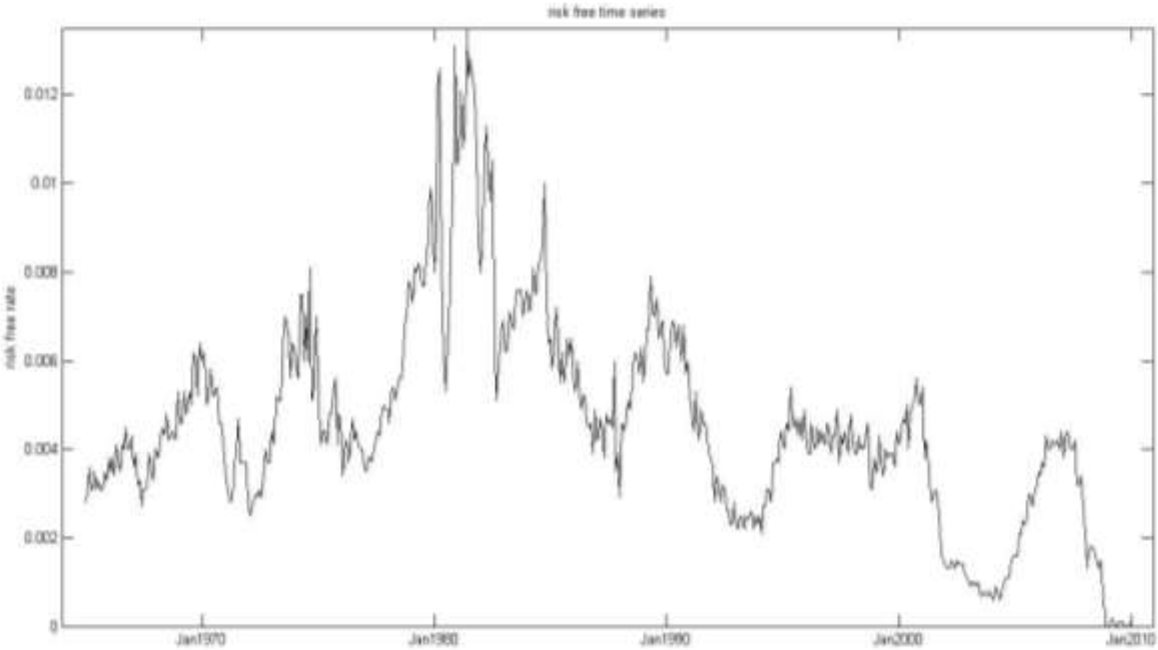
Figure 6 tbond cumulative return time series plot



From the two plots above it is immediately apparent that the time series shows two distinct behaviors over the sample period. In the first period, namely between January 1965 and the first years of the 1980s, an overall slight decrease occurred; in the second period a consistent uptrend occurred throughout the entire period, it then reached a peak of about 2 almost by the end of the period.

The 1-month US T-bill rate, hereafter called risk_free, has showed, over the sample period, an average value equal to 0.46%, a standard deviation of 0.23%, which is equivalent to a variance of about 0.0001%. Since the risk free rate is not one of the asset classes the investor is assumed to invest in, the Sharpe ratio has not been computed. The median is equal to 0.43% and the maximum and minimum sample values are respectively 1.35% and 0%. In addition a normal distribution has been estimated from the sample data, the estimation returns an estimate of the mean equal to 0.46%, which is equivalent to an annualized mean of 5.47%, and an estimate of the standard deviation of 0.23%, which is equivalent to an annualized standard deviation of 0.81%. The time series exhibits a substantial positive skewness equal to 0.8436 and a kurtosis in excess of the Gaussian benchmark (equal to three) equal to 4.5912. The risk_free time series is depicted in the plot below.

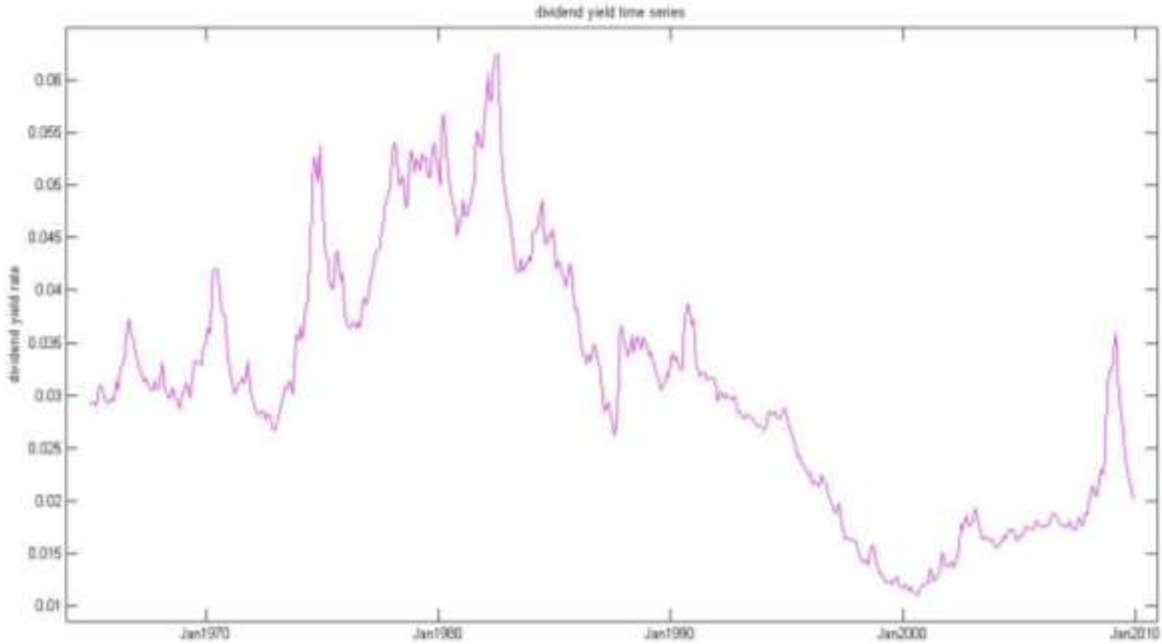
Figure 7 risk_free time series plot



From the risk_free plot it can be seen that the 1-month US T-Bill rate has been characterized by significant fluctuations; however, two distinct overall trends can be easily detected. An uptrend occurred in the first period and culminated in the first years of the 1980s at nearly 1.4% followed by a downtrend which reached its minimum at nearly 0% by the end of the sample period.

The S&P500 dividend yield, hereafter called div_y, has showed, over the sample period, an average value equal to 3.11%, a standard deviation of 1.20%, which is equivalent to a variance of about 0.01%. Since the risk free rate is not one of the asset classes the investor is assumed to invest in, the Sharpe ratio has not been computed. The median is equal to 3.08% and the maximum and minimum sample values are respectively 1.11% and 6,24%. In addition a normal distribution has been estimated from the sample data, the estimation returns an estimate of the mean equal to 3.11% and an estimate of the standard deviation equal to 0.23%. Since the time series already represents an annual measure, i.e. the annual dividends paid by the companies in the S&P500 index divided by the price index, the estimated mean and standard deviation are not further annualized. The time series exhibits a substantial positive skewness equal to 0.8436 and a kurtosis in excess of the Gaussian benchmark (equal to three) equal to 4.5912. The div_y time series is depicted in the plot below.

Figure 8 div_y time series plot



The figure above shows the evolution of the S&P500 dividend yield over the sample period. The div_y time series characterized by a first period of relatively high volatility followed by a less volatile one. Even if in the first period the dividend yield fluctuates significantly, a general uptrend that culminated just above 6% in the first years of 1980s can be detected. Since the mid-1980s a steep downtrend happened, the period ended with a fall to nearly 0% by the end of the 2000s followed by a considerable rise in the last decade of the sample.

2.4 A comparative analysis of the asset classes returns

In this paragraph a comparison of the three asset class is provided with the aim to detect the similarities and differences between them, to identify peculiarities that make each one of the time series unique, and to study how they comove over the entire sample period. In doing so some figures and plots are provided with the intent to make the analysis more clear and comprehensible.

Figure 9 lo20, hi20 and tbond time series plot

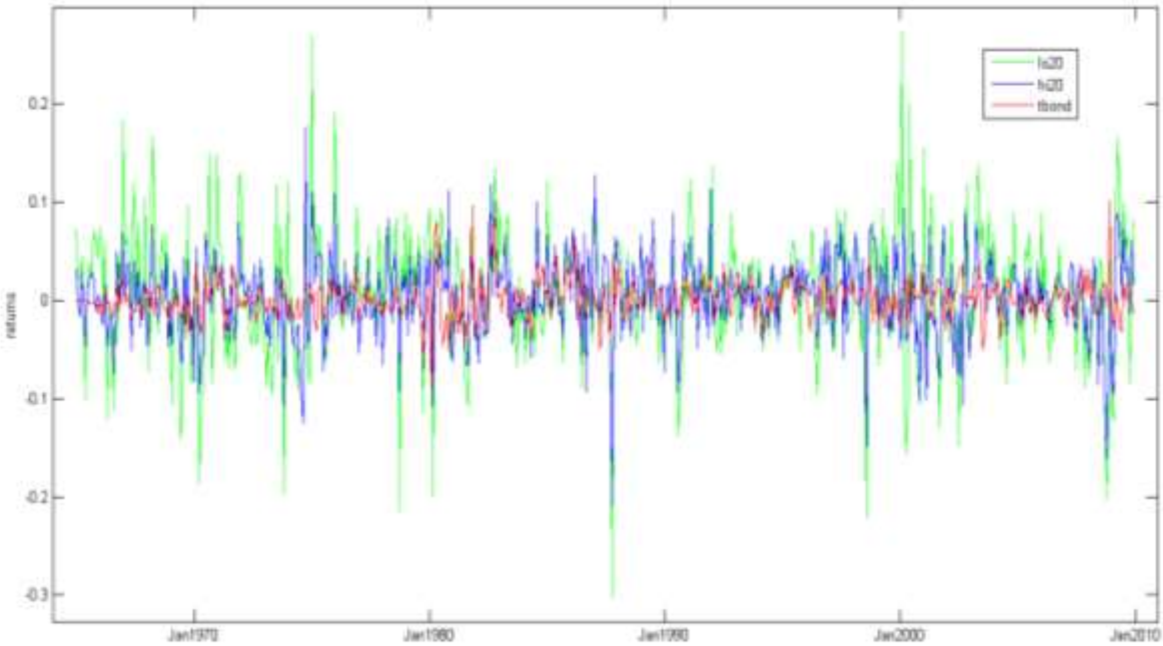


Figure 10 lo20, hi20 and tbond cumulative returns time series plot

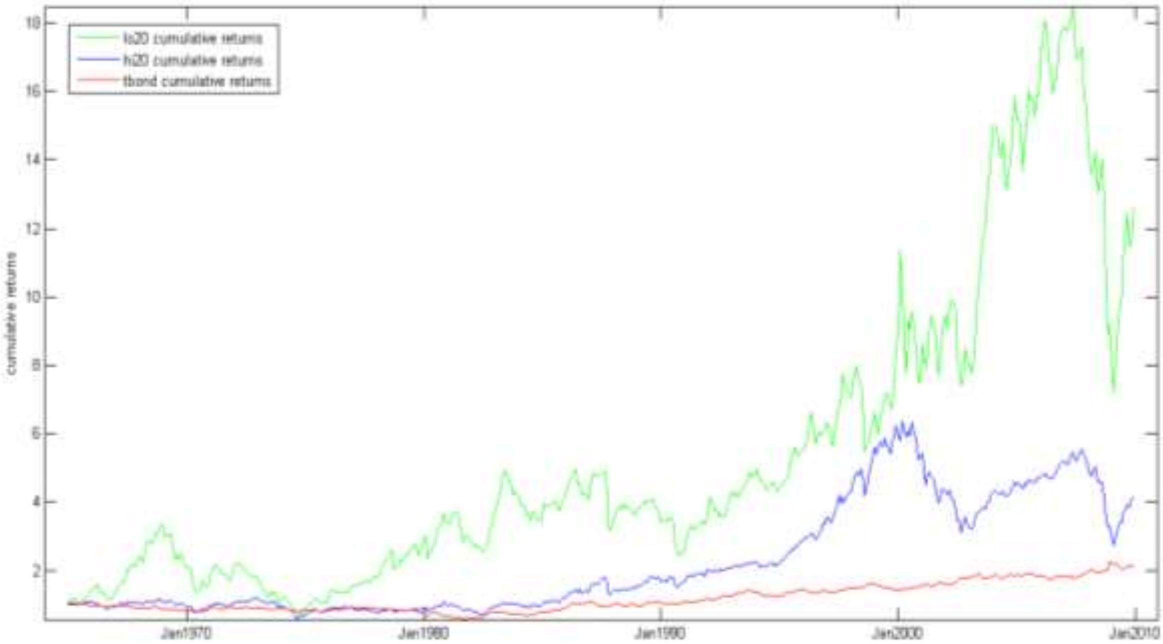


Figure 11 lo20, hi20 and tbond scatter plots and histogram matrix

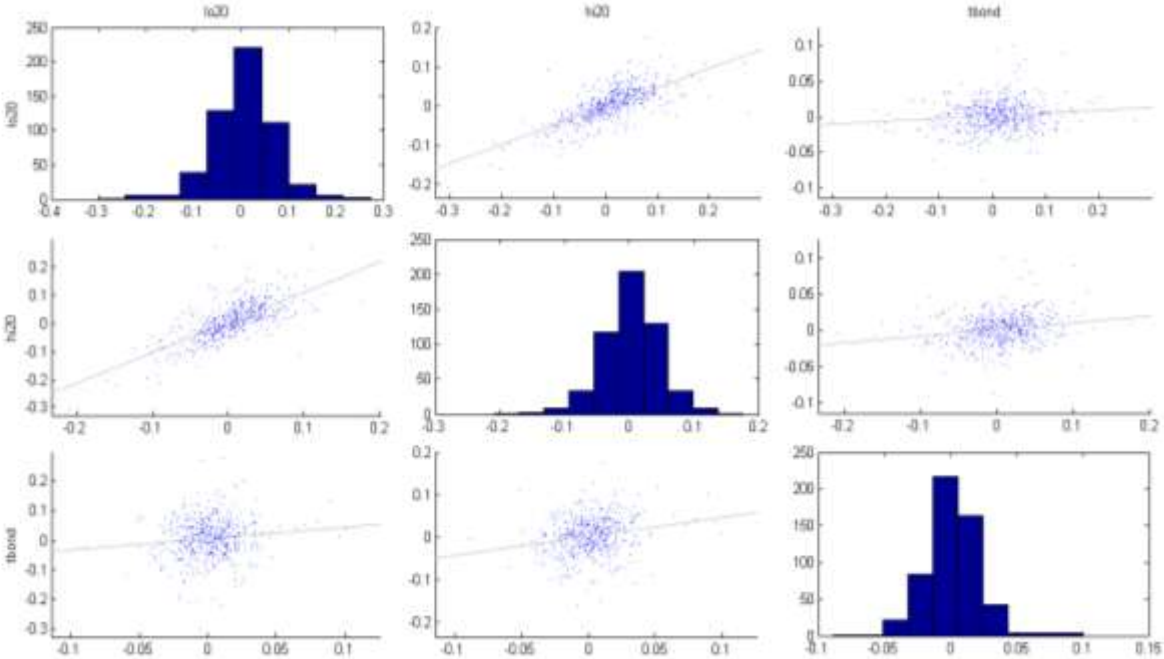


Figure 9 shows the three asset class time series of returns in the same plot; from this figure it can be recognized that the lo20 time series, over the entire sample period, has experienced wider fluctuations than hi20 and tbond with the latter one characterized to be the least volatile asset class. It can be also noted that some volatility clusters occurred in a pretty synchronized way among the asset class, as shown by the data between 1995 and 2005; this feature of the data may lead to consider that the three asset class have been positively correlated and perhaps caused by a common exogenous phenomena. Figure 10 shows the evolution of the accumulated returns of one unit of wealth invested in each one of the three asset classes. As widely suggested by the literature the US small cap generally yields a higher average return and a higher volatility, this peculiarity is clearly witnessed by the sample data, in fact the lo20 time series reacted significantly faster and sharper than hi20 and tbond after swing points in the data, leading to a higher standard deviation. The evolution of the hi20 cumulative returns shows a trend generally comparable to the one of lo20 with the difference that the former one has been characterized to be much less volatile with a period of substantial lack of fluctuations that is extended through the sample data from the 1965 to the 1990. The tbond cumulated returns exhibits a slight and constant uptrend that had

reached its highest point just above 2 (100% cumulated returns) by the end of the sample data period. The almost total absence of downtrend phases, opposed to the wide fluctuations that have been experienced by the two equity asset classes, may suggest that the tbond time series is uncorrelated, or at least very slightly correlated, with the two equity asset classes time series. From the examination of the two equity time series, especially lo20, is straightforward to identify the moments in which the financial crises occurred; a first downtrend started at the end of the 1980s coinciding with the 1987 Black Monday, the largest one-day percentage decline in stock market history; a second significant downtrend started around 2000 following the concern about the internet companies (perhaps many of them were small cap companies and then comprised in the first and second CRSP size decile US equity portfolios and thereafter represented by the lo20 time series); a third downtrend started approximately at 2007 coinciding with the Global Financial Crises. The fact that the sample extends well into 2008-2009 allows me to reach the conclusion that it is fully affected by the recent turmoil in international equity markets. Additionally, the bull and bear phases present in the data lead to the conclusions that the selected sample period is probably representative of market regimes. All the three asset classes time series are characterized by a kurtosis in excess of the Gaussian benchmark (three). However, only the tbond time series has a positive skewness coefficient, while both lo20 and hi20 have negative skewness coefficients. Figure 11 consists in a matrix of subaxes containing scatter plots of the asset class in the vertical axis against the asset class in the horizontal axis, the diagonal are replaced with histogram plots of the asset class in the corresponding position along the vertical or horizontal axis. The trend line superimposed on each scatter gives the direction of the correlation between two time series, while the closer the dots are to this line, the stronger the correlation between the time series. Although all the linear correlation coefficient between the three asset classes are positive, some of them are higher than others. The strength of the linear correlation between lo20 and hi20 is the highest and is equal to 0.7180, the one between lo20 and tbond is the lowest and is equal to 0.1210 while the one between hi20 and tbond is equal to 0.2171.

2.5 Stationarity and normality tests

In this paragraph the results of the normality and stationarity tests are provided and described for the three asset classes time series and the dividend yield. Additional figures are provided with the intent to make the dissertation clearer and more comprehensive.

Firstly, I started by performing normality test and secondly stationarity test. In the attempt to determine if each time series was well-modeled by a normal distribution I proceeded adopting in the very first stage an informal graphical method that consisted in comparing a histogram of the sample data to a Normal probability curve. I further examined the time series with the help of the normal probability plot, the quantile-quantile plot (QQ plot), the kernel smoothing probability density estimate and the empirical cumulative distribution function for each time series. In the second stage I performed some common statistical hypothesis tests in which data are tested against the null hypothesis that they are normally distributed. In the first part of the next passage the results of the graphical method is provided for each time series while at the end of the passage the results of the hypothesis test, contained in a table, are exposed.

The interpretation of the histogram and the superimposed normal density distribution fitted to the data is the following: if the time series has been drawn from a Normal distribution then empirical distribution of the data, represented by the histogram, must be bell-shaped and resemble the normal distribution fitted, otherwise a suspect that the data are not normally distributed should rise. The normal probability plot has the sample data displayed with the plot symbol '+'. Superimposed on the plot is a line joining the first and third quartiles of the sample data. This line is extrapolated out to the ends of the sample to help evaluate the linearity of the data. The purpose of a normal probability plot is to graphically assess whether the sample data could come from a normal distribution. If the data are normal the plot will be linear. Other distribution types will introduce curvature in the plot. The QQ displays a quantile-quantile plot of the quantiles of the sample data versus theoretical quantiles from a normal distribution. If the distribution of the sample data is normal, the plot will be close to linear. The plot has the sample data displayed with the plot symbol '+'. Superimposed on the plot is a line joining the

first and third quartiles of the sample data. This line is extrapolated out to the ends of the sample to help evaluate the linearity of the data. The kernel smoothing probability density estimate returns a probability density estimate for the sample data. The estimate is based on a normal kernel function, and is evaluated at 100 equally spaced points, that cover the range of the sample data. If the sample data is normally distributed then the plot of the kernel smoothing probability density estimate must resemble the shape of a normal density curve, otherwise a suspect that the data are not normally distributed should rise.

Figure 12 lo20 frequency histogram and normal density function fit

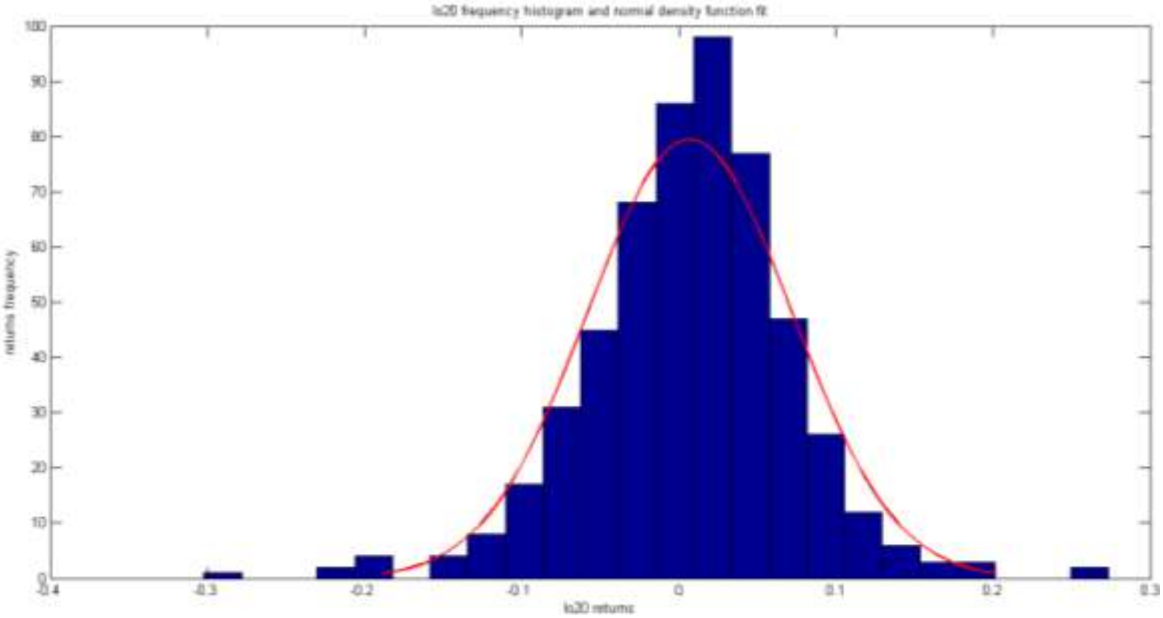


Figure 13 lo20 normal probability plot

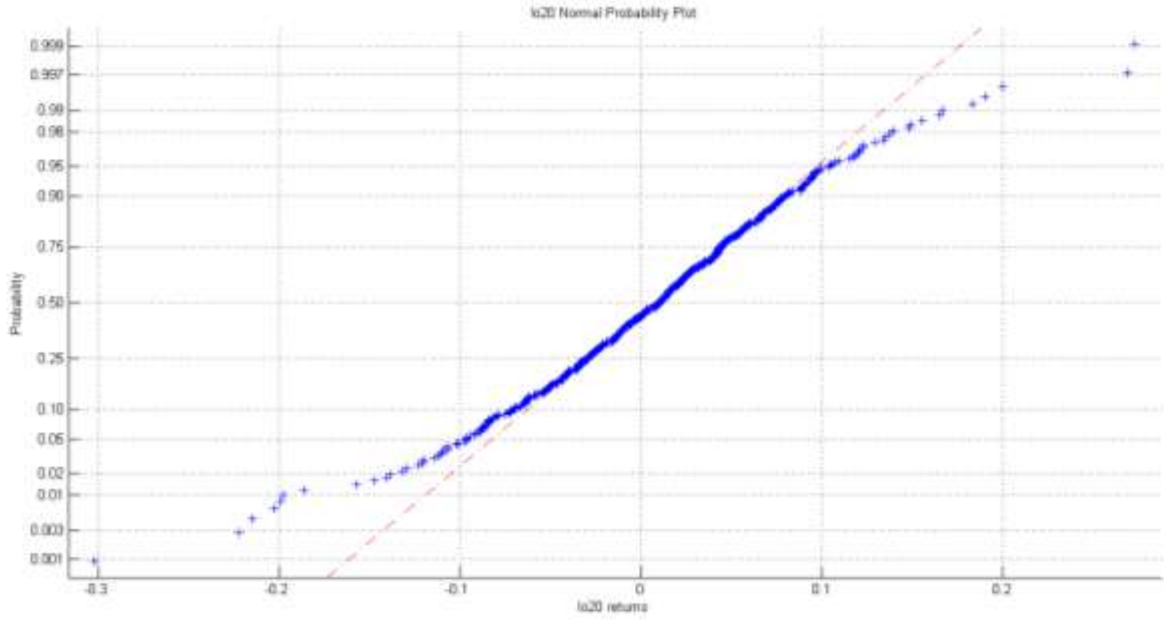


Figure 14 lo20 QQ plot

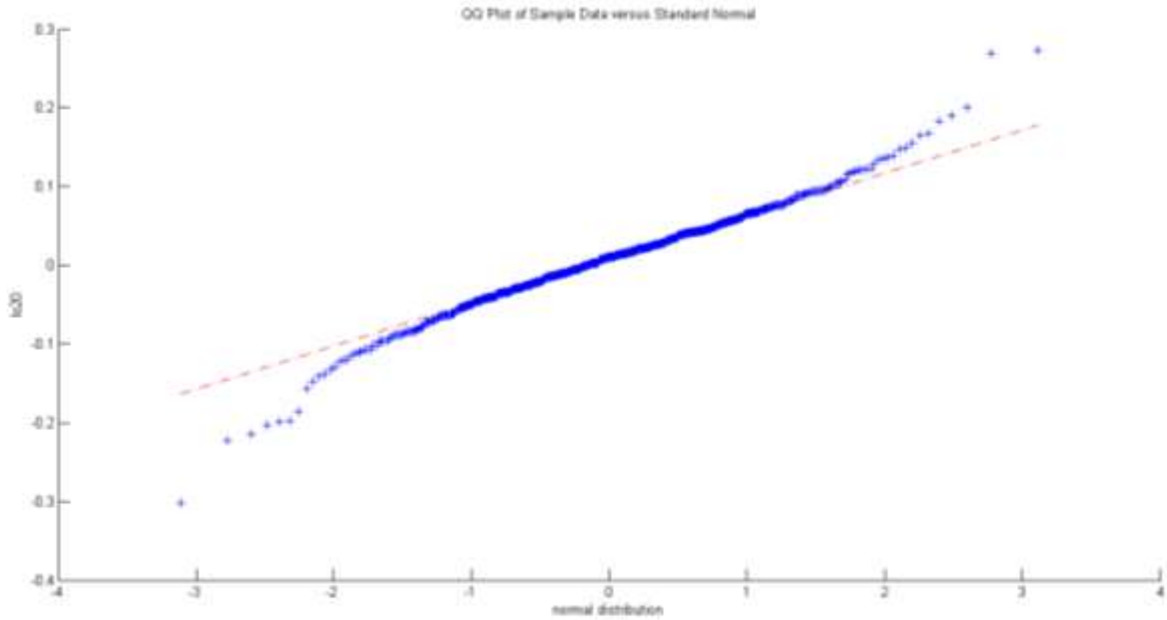


Figure 15 lo20 kernel smoothing probability density estimate

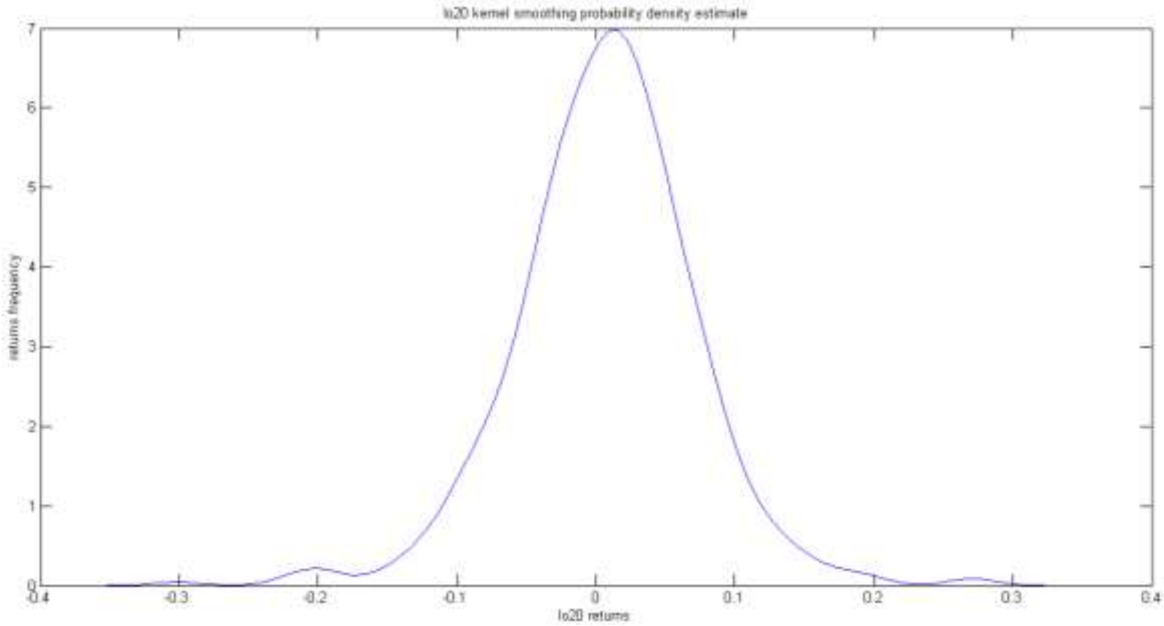
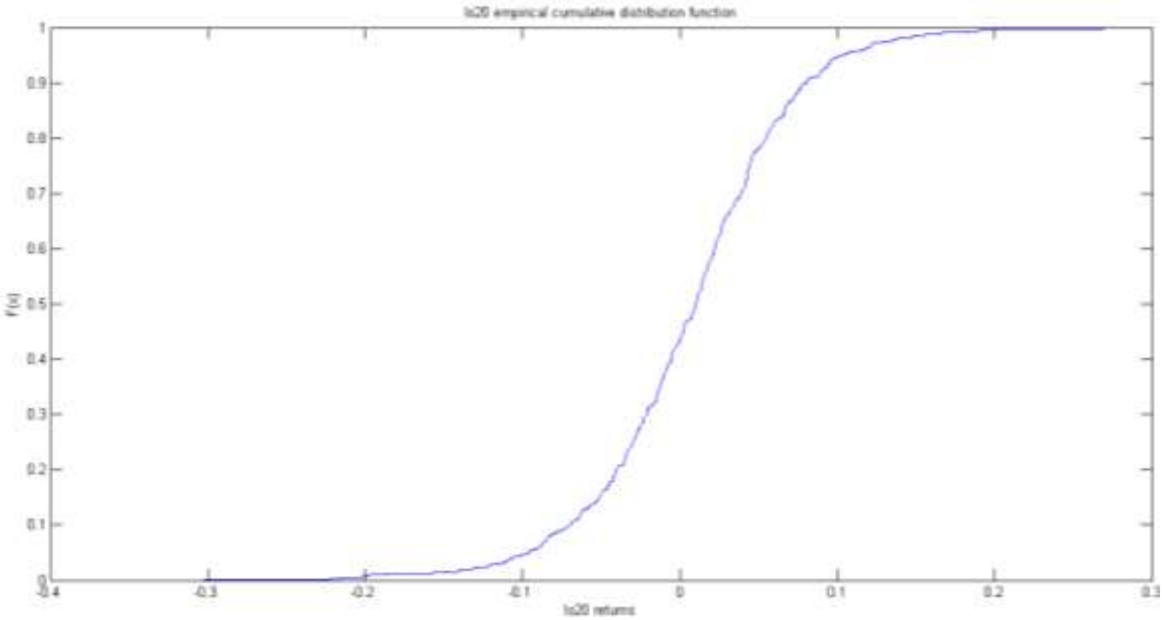


Figure 16 lo20 empirical cumulative distribution function



From the examination of the five plots above the lo20 time series clearly appeared to be not normally distributed. Figure 12 shows that the lo20 probability density distribution presents fat tails and, returns close to zero have a significant higher probability of happen than in the case of a normal distribution. Figure 13 and 14 both shows the presence of outliers returns that diverge significantly from the straight line that join the first and the third quantiles. Figure 15 points out the

presence of the fat tail characteristic represented by the fact that the kernel smoothing probability density estimate does not appear to have the shape of a normal distribution, on the contrary it shows returns far from the mean higher density than the normal distribution.

Figure 17 hi20 frequency histogram and normal density function fit

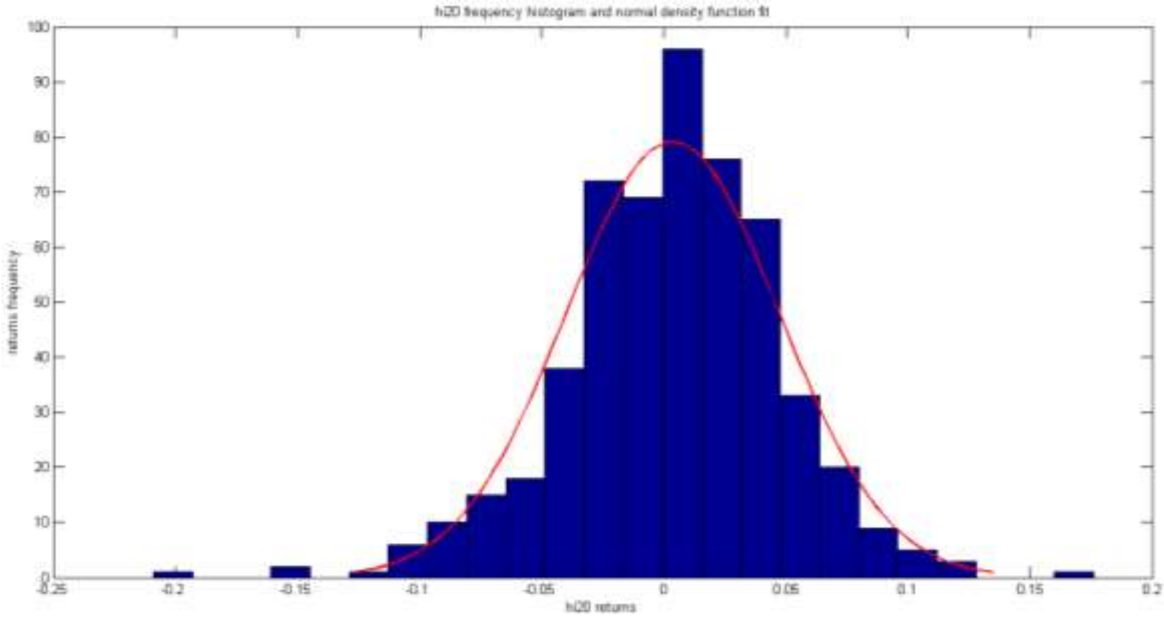


Figure 18 hi20 normal probability plot

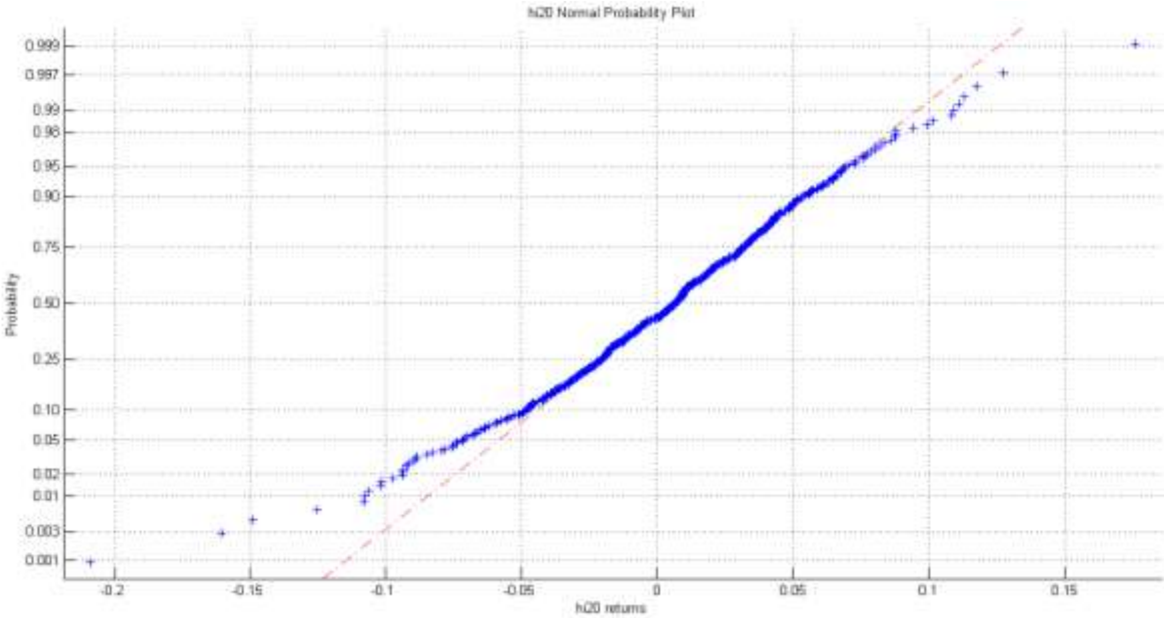


Figure 19 hi20 QQ plot

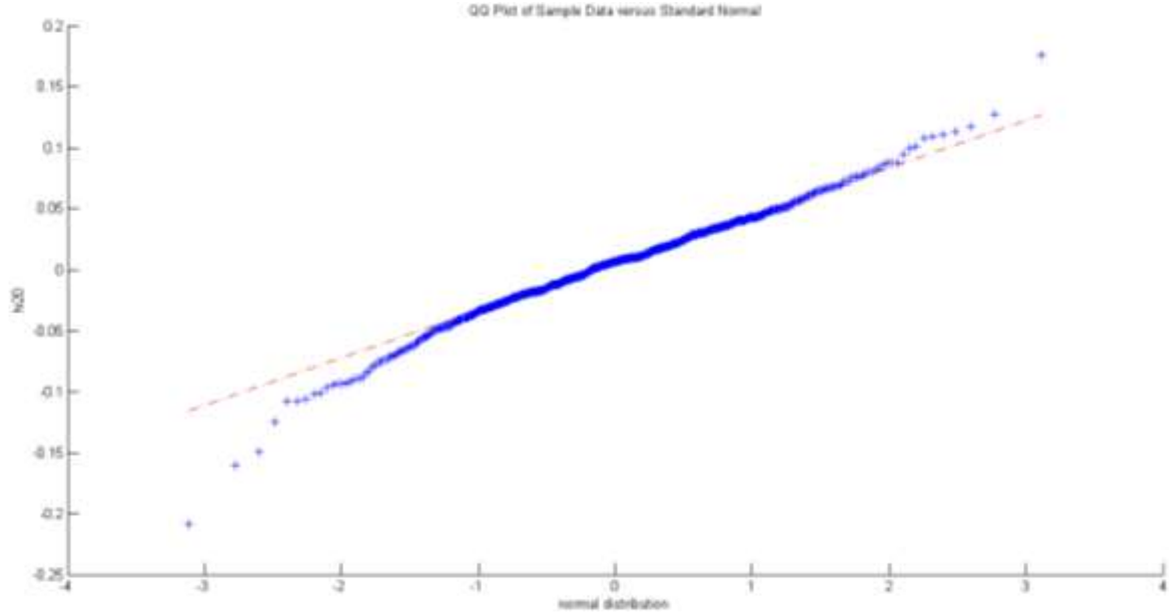


Figure 20 hi20 kernel smoothing probability density estimate

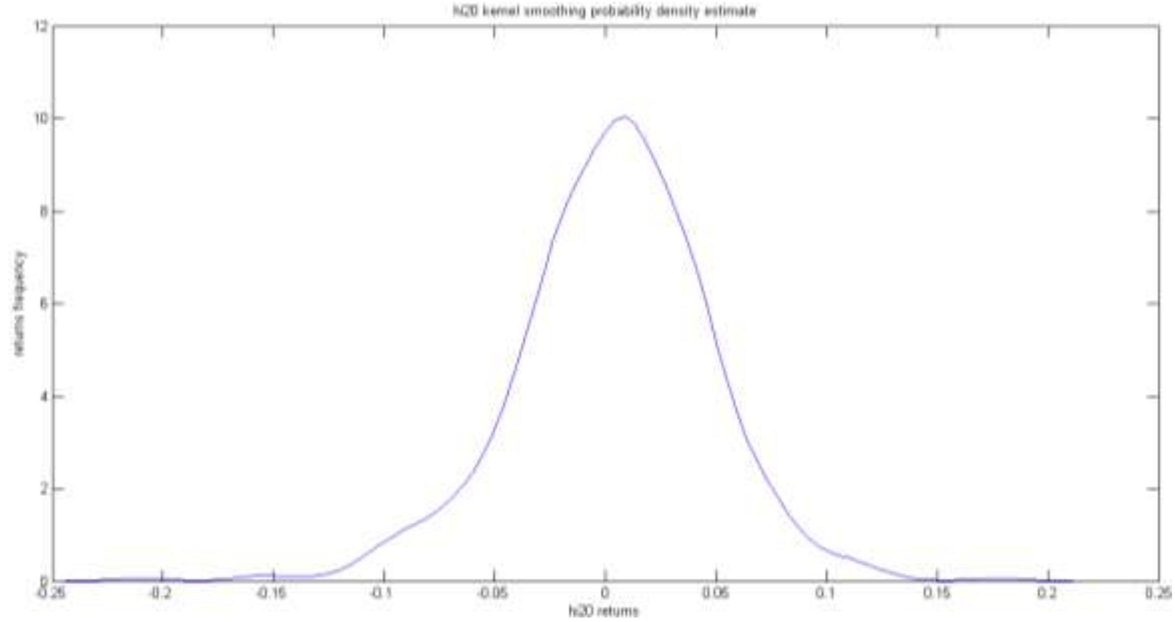
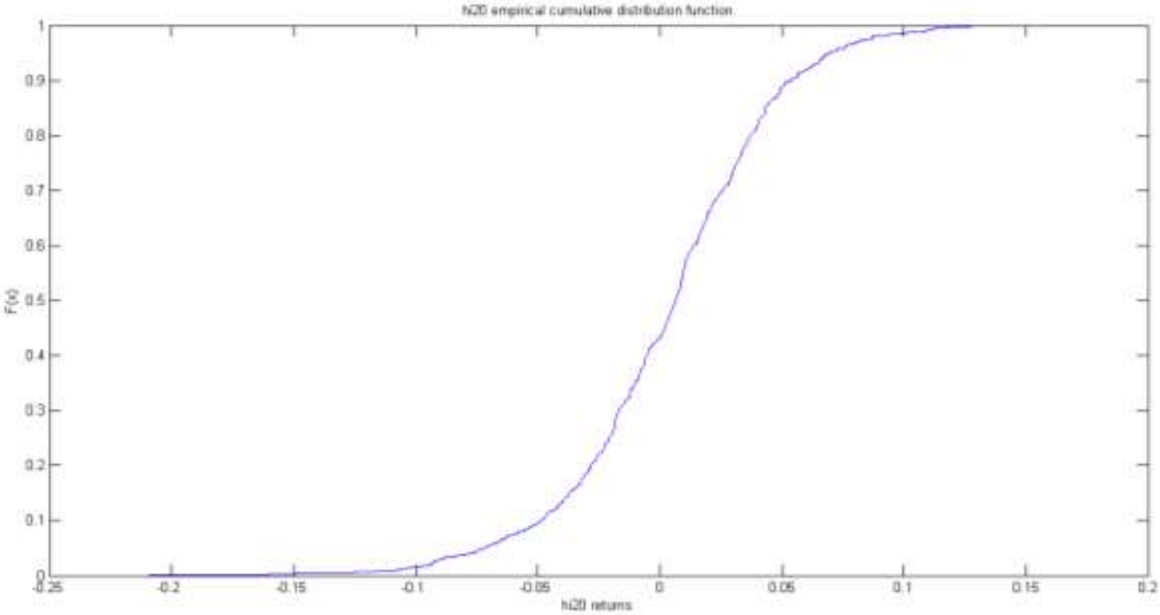


Figure 21 hi20 empirical cumulative distribution function



The hi20 time series at a first glance seems to be very similar to lo20 regarding the distribution characteristic. From the examination of the five plots above the hi20 time series clearly appeared to be not normally distributed. Figure 17 shows that the hi20 probability density distribution presents fat tails and, returns close to zero have a significant higher probability of happen than in the case of a normal distribution. Figure 18 and 19 both show the presence of outlier returns that diverge significantly from the straight line that join the first and the third quantiles. Figure 20 points out the presence of the fat tail characteristic represented by the fact that the kernel smoothing probability density estimate does not appear to have the shape of a normal distribution, on the contrary it shows, for returns far from the mean, higher density than the normal distribution.

Figure 22 tbond frequency histogram and normal density function fit

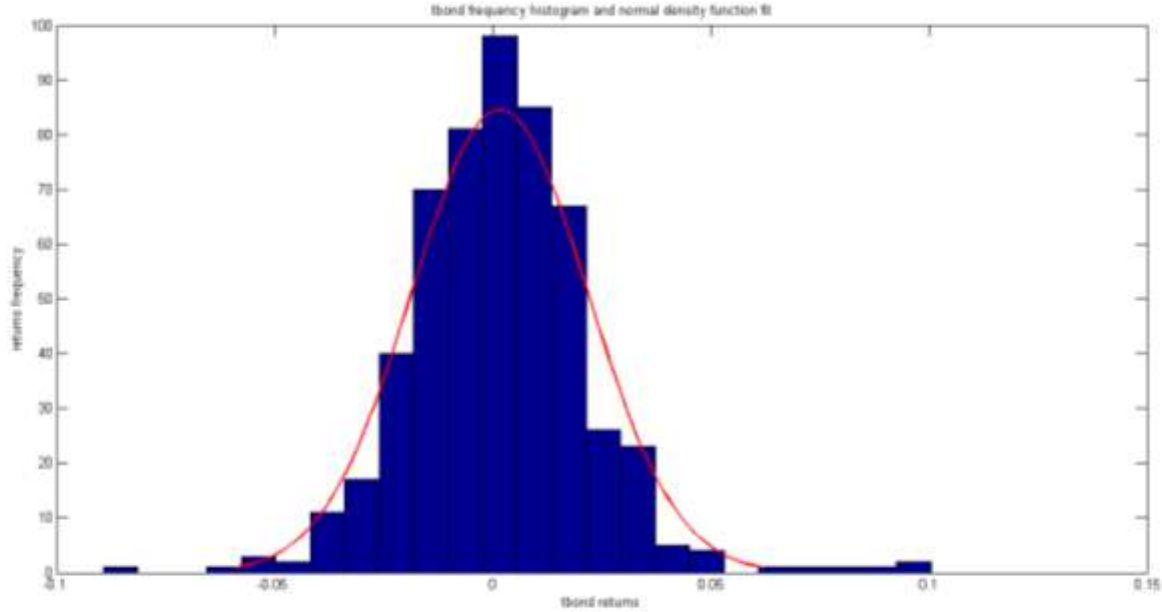


Figure 23 tbond normal probability plot

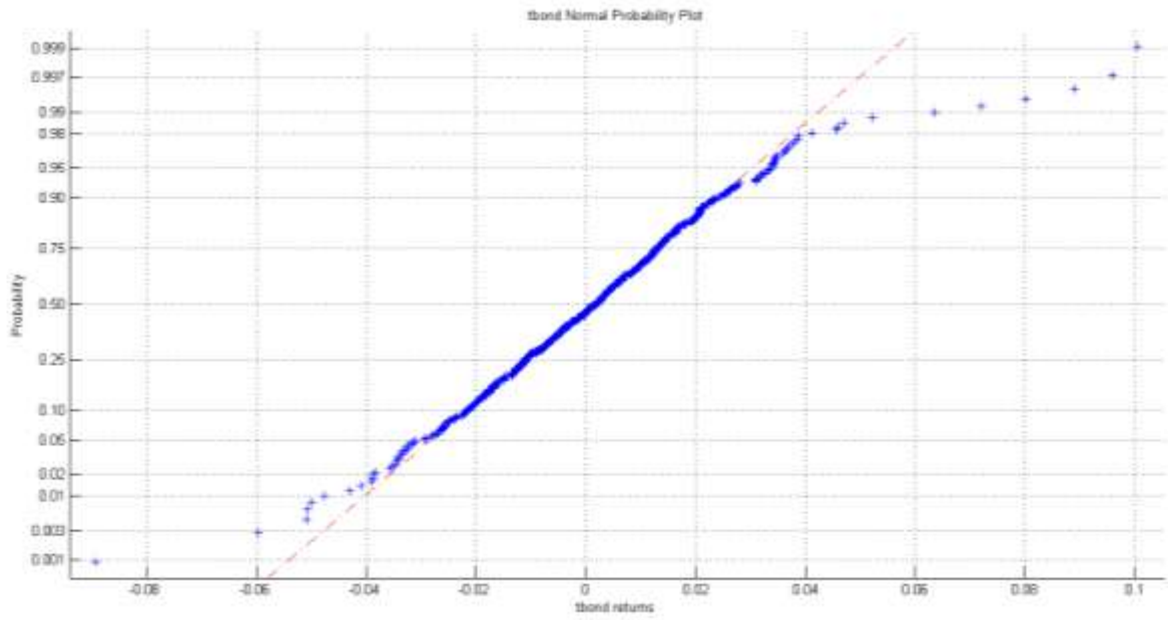


Figure 24 tbond QQ plot

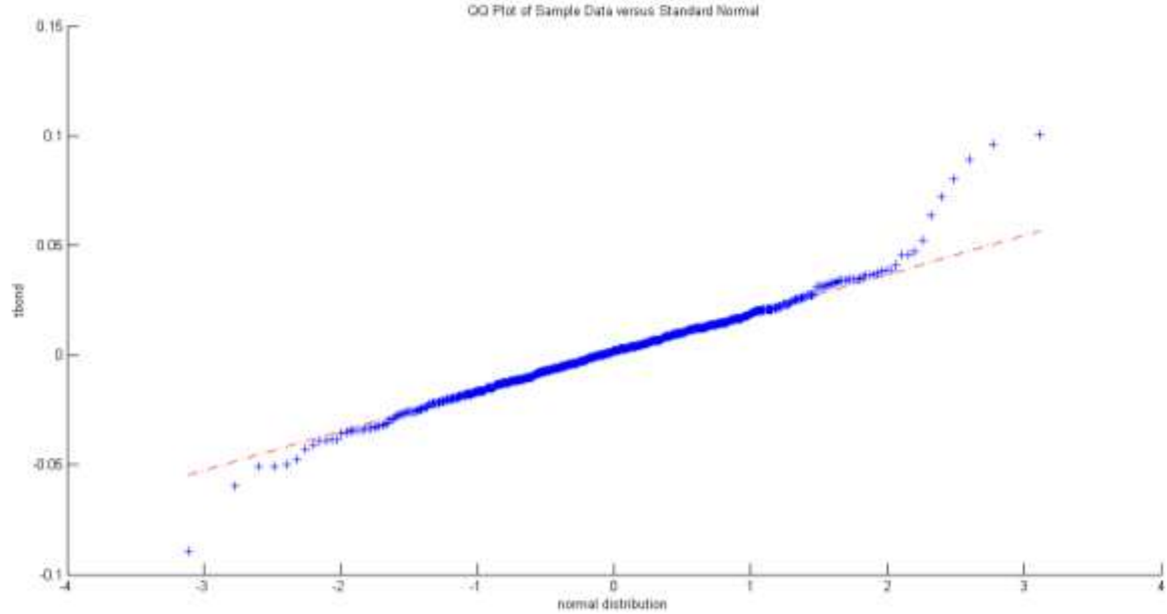


Figure 25 tbond kernel smoothing probability density estimate

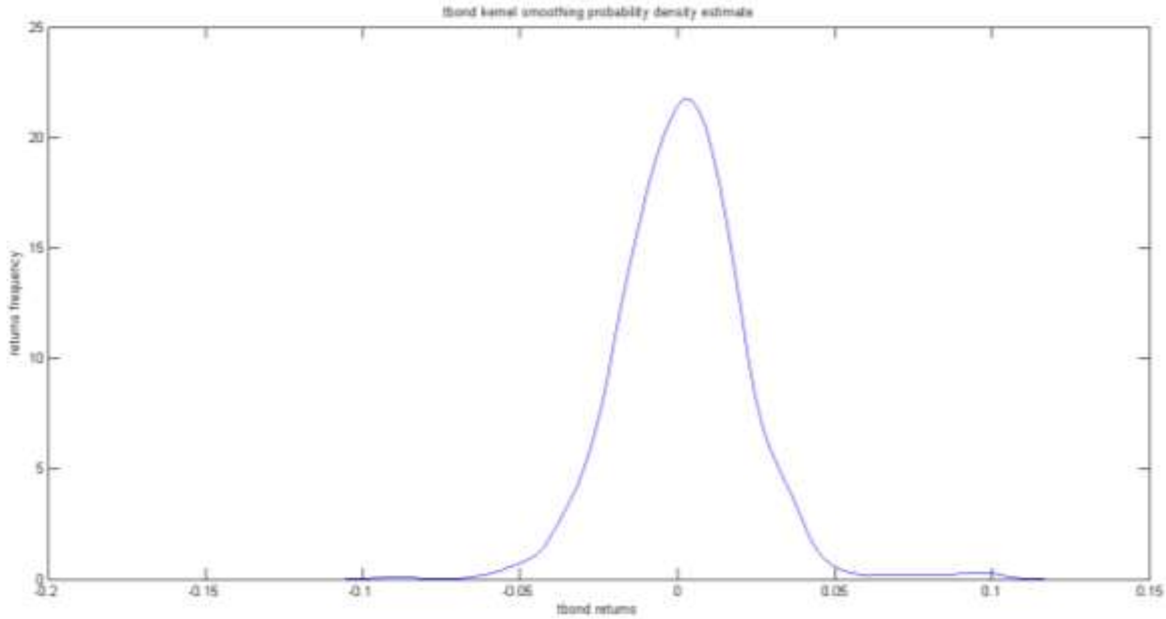
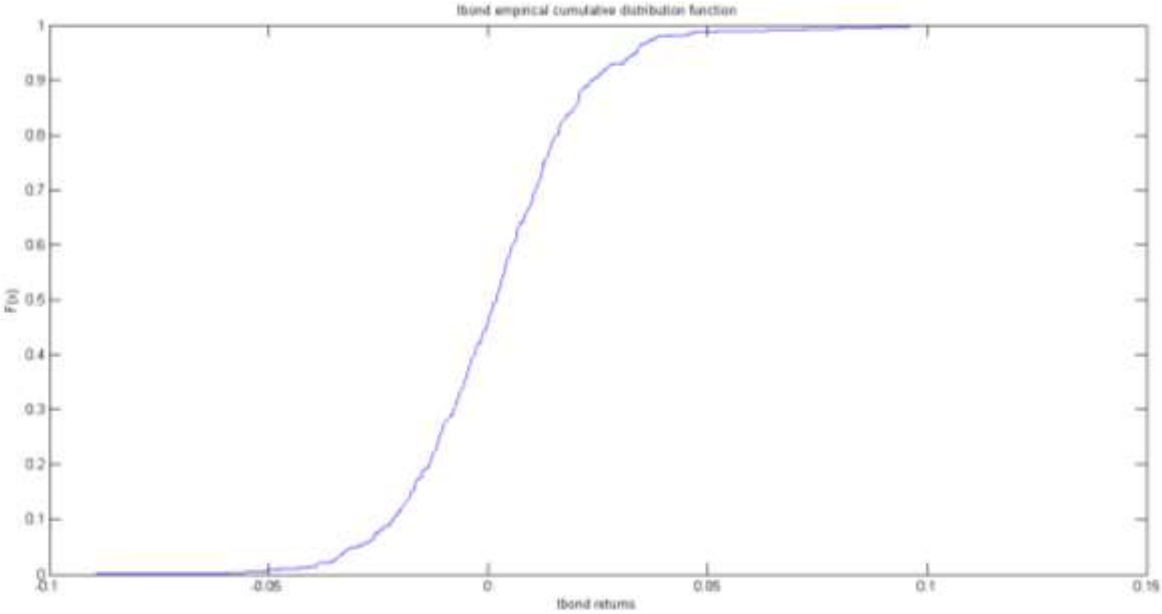


Figure 26 tbond empirical cumulative distribution function



From the examination of the five tbond plots above emerges that the time series is characterized to have a significant positive skewness that may prevent it to be normally distributed. Figure 22 shows that the tbond probability density distribution presents fat tails, large skewness and, returns close to zero have a slight higher probability of happen than in the case of a normal distribution. Figure 23 and 24 both shows the presence of outlier returns that diverge significantly from the straight line that join the first and the third quantiles. Figure 25 points out the presence of a large positive skewness by the fact that the kernel smoothing probability density estimate does not appear to have the shape of a normal distribution, on the contrary it shows a marked asymmetry.

Figure 27 div_y frequency histogram and normal density function fit

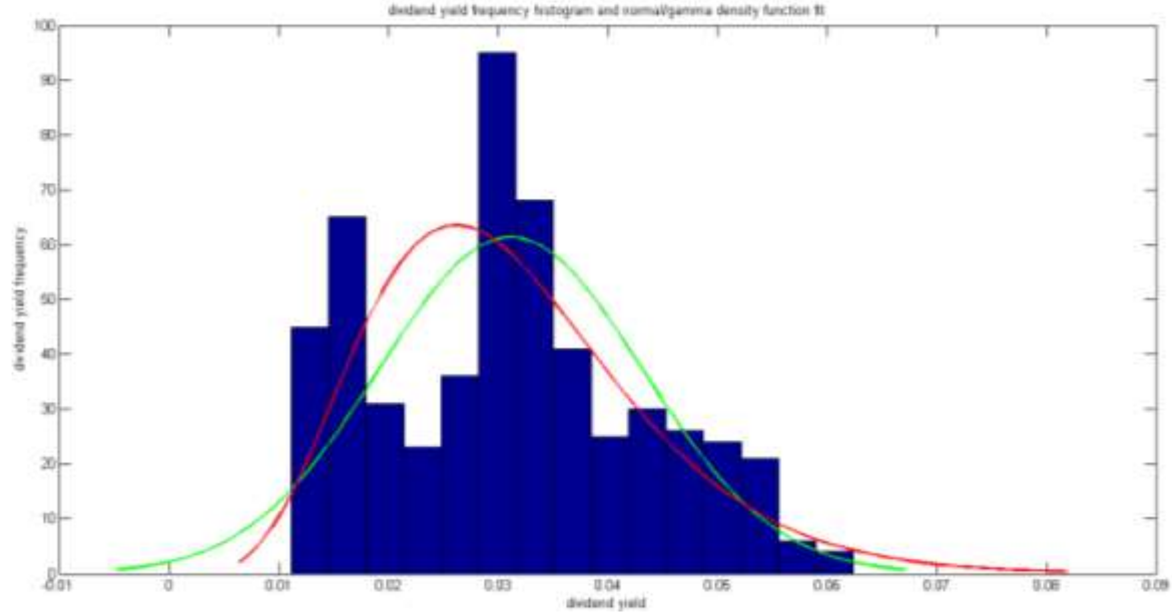


Figure 28 div_y normal probability plot

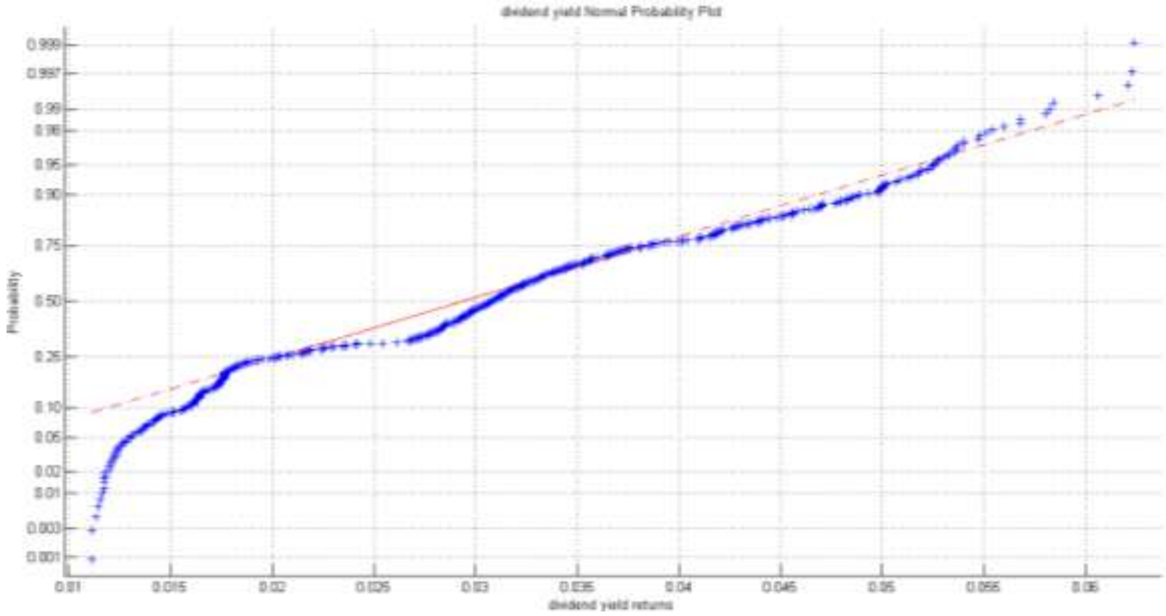


Figure 29 div_y QQ plot

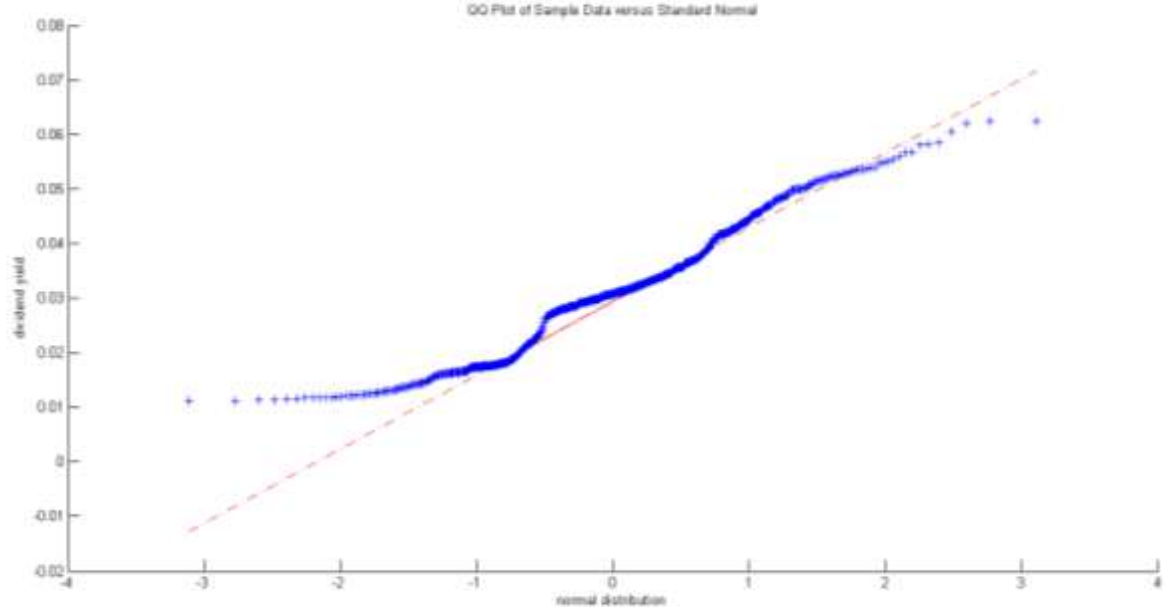


Figure 30 div_y kernel smoothing probability density estimate

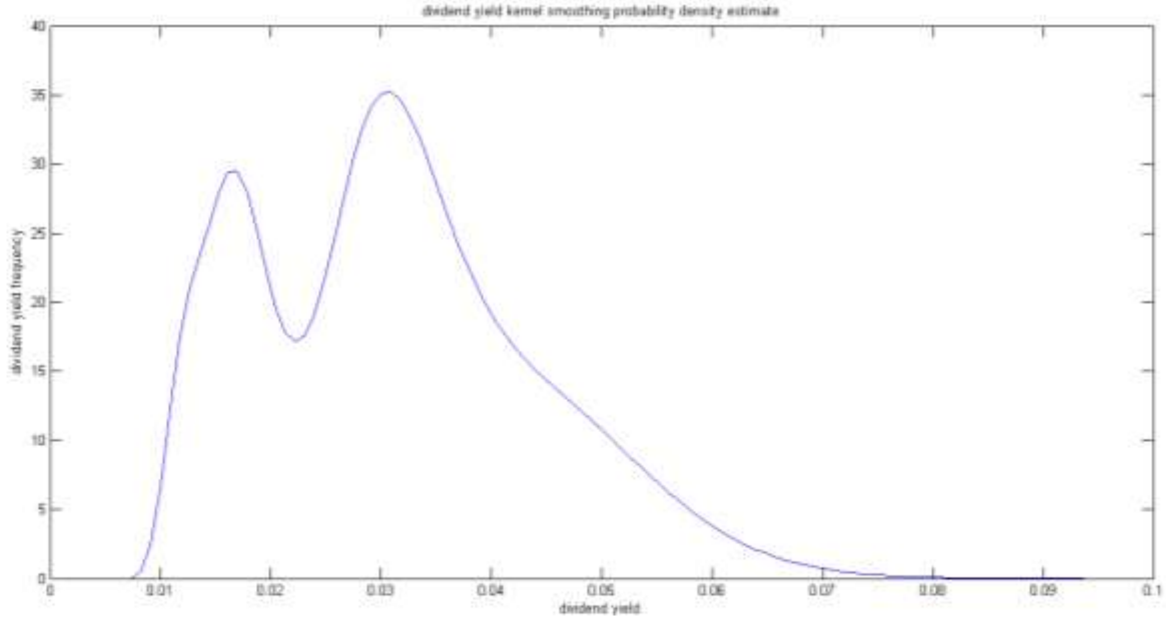
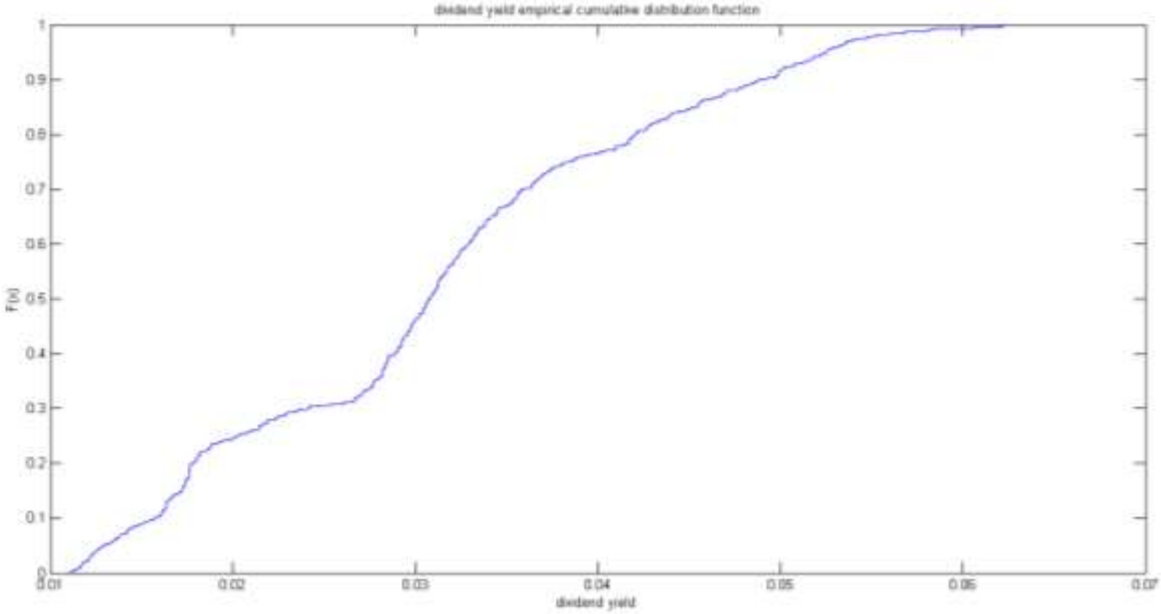


Figure 31 div_y empirical cumulative distribution function



From the examination of the five plots above the div_y time series appeared extremely not normally distributed. In Figure 27 I decided to show, in addition to a fitted normal distribution (green solid line), a gamma distribution (red solid line) which has in common with the div_y sample a positive probability density support. From a graphical comparison of the degree of fit of the two different fitted distributions to the sample date appears clear that the div_y is not normally distributed. Figure 28 and 28 both shows the presence of outlier returns that diverge significantly from the straight line that join the first and the third quantiles. Figure 30 points out also the fact that the probability density distribution of the sample data is even not unimodal, which is in turn a large divergence from the normal distribution case.

Even though at this point of the analysis it appeared quite likely that all the time series were not normally distributed I further analyzed them performing some common statistical hypothesis tests in which data are tested against the null hypothesis that they are normally distributed. In the following table the results are exhibited.

Table 3 Normality Tests Results

test name	alpha	object	lo20	hi20	div_y	tbond

Chi-square goodness-of-fit	10%	decision	1	1	1	1
Chi-square goodness-of-fit	10%	p-value	0.0014	0.0348	0	0.012
Chi-square goodness-of-fit	5%	decision	1	1	1	1
Chi-square goodness-of-fit	5%	p-value	0.0014	0.0348	0	0.012
Chi-square goodness-of-fit	1%	decision	1	0	1	0
Chi-square goodness-of-fit	1%	p-value	0.0014	0.0348	0	0.012
Anderson-Darling	10%	decision	1	1	1	1
Anderson-Darling	10%	p-value	0	0	0	0
Anderson-Darling	10%	statistic	2.28	1.78	4.81	2.14
Anderson-Darling	10%	cvalue	0.62	0.62	0.61	0.65
Anderson-Darling	5%	decision	1	1	1	1
Anderson-Darling	5%	p-value	0	0	0	0
Anderson-Darling	5%	statistic	2.28	1.78	4.81	2.14
Anderson-Darling	5%	cvalue	0.80	0.74	0.74	0.75
Anderson-Darling	1%	decision	1	1	1	1
Anderson-Darling	1%	p-value	0	0	0	0
Anderson-Darling	1%	statistic	2.28	1.78	4.81	2.14
Anderson-Darling	1%	cvalue	0.94	0.97	1.13	1.04
Jarque-Bera	10%	decision	1	1	1	1
Jarque-Bera	10%	p-value	0	0	0.0025	0
Jarque-Bera	10%	statistic	119.64	78.20	16.63	246.75
Jarque-Bera	10%	cvalue	4.36	4.34	4.35	4.37
Jarque-Bera	5%	decision	1	1	1	1
Jarque-Bera	5%	p-value	0	0	0	0
Jarque-Bera	5%	statistic	119.64	78.20	16.63	246.75
Jarque-Bera	5%	cvalue	5.87	5.89	5.85	5.84
Jarque-Bera	1%	decision	1	1	1	1
Jarque-Bera	1%	p-value	0	0	0	0
Jarque-Bera	1%	statistic	119.64	78.20	16.63	246.75
Jarque-Bera	1%	cvalue	10.65	10.82	10.69	10.66

Three different normality tests have been performed in the attempt to assess how likely each one of the three asset class time series of returns, individually considered, could have been drawn from a normal density distribution. The Chi-square goodness-of-fit is a test that returns a decision for the null hypothesis that the sample data comes from a normal distribution with a mean and variance estimated from the sample data, using the chi-square goodness-of-fit test. The alternative hypothesis is that the data does not come from such a distribution. The result (the row “decision” in Table 3) is 1 if the test rejects the null hypothesis at

the given significance level, and 0 otherwise. The test groups the data into bins, calculating the observed and expected counts (based on the hypothesized distribution) for those bins, and computing the chi-square test statistic. The test statistic has an approximate chi-square distribution when the counts are sufficiently large. The Andreson-Darling is a test that returns a decision for the null hypothesis that the sample data comes from a population with a normal distribution, using the Anderson-Darling test. The alternative hypothesis is that the sample data does not come from a population with a normal distribution. The result (the row “decision” in Table 3) is 1 if the test rejects the null hypothesis at the given significance level, or 0 otherwise. The test statistic belongs to the family of quadratic empirical distribution function statistics, which measure the distance between the hypothesized distribution and the empirical one. The weight function for the Anderson-Darling test places greater weight on the observations in the tails of the distribution, thus making the test more sensitive to outliers and better at detecting departure from normality in the tails of the distribution. The Jarque-Bera test returns a test decision for the null hypothesis that the sample data comes from a normal distribution with an unknown mean and variance. The alternative hypothesis is that it does not come from such a distribution. The result (the row “decision” in Table 3) is 1 if the test rejects the null hypothesis at the given significance level, and 0 otherwise. The Jarque–Bera test is a goodness-of-fit test of whether sample data have the skewness and kurtosis matching a normal distribution. If the data come from a normal distribution, the JB statistic asymptotically has a chi-squared distribution with two degrees of freedom, so the statistic can be used to test the hypothesis that the data are from a normal distribution. The null hypothesis, in other words, is a joint hypothesis of the skewness being zero and the excess kurtosis being zero. The three tests returns for any significance level the decision that the four time series are statistically not normally distributed except for the Chi-square goodness-of-fit that failed to reject the null hypothesis at 1% significance level for ho20 and tbond.

From the point of view of this dissertation, it is important to note that all the time series displayed significant deviation from the normal distribution benchmark, as evidenced by the statistically significant normality tests results, and then they can

be considered not normally distributed. This is a clear indication that the returns on the three asset classes cannot be captured by linear models to reinforce the idea of a regime switching model, which is more flexible in accommodating the mixing of several empirical distributions. The regime switching model is also supported by the results of normality tests, as all null hypotheses are strongly rejected. Since regime switching models account for non-normality by using a mixture-of-normal distributions approach, they deliver a more accurate way of modeling the dynamics and the distribution of the three asset classes returns than models using only one normal distribution.

In the following part of the paragraph stationarity tests results are provided and commented. These tests can be used to determine if trending data should be first differenced or regressed on deterministic functions of time to render the data stationary. A process is said to be covariance-stationary, or weakly stationary, if its first and second moments (hence also the covariance) are time invariant. In other words the structure of the series does not change with the time. A stationary series is relatively easy to predict since its statistical properties will be the same in the future as they have been in the past.

Table 4 Stationarity Tests Results

test name	**alpha**	**object**	**lo20**	**hi20**	**div_y**	**tbond**
Augmented Dickey-Fuller TS LAG 0	10%	decision	1	1	0	1
Augmented Dickey-Fuller TS LAG 0	10%	p-value	0.00	0.00	0.60	0.00
Augmented Dickey-Fuller TS LAG 0	10%	statistic	-18.29	-22.08	-1.98	-16.93
Augmented Dickey-Fuller TS LAG 0	10%	cvalue	-3.13	-3.13	-3.13	-3.13
Augmented Dickey-Fuller TS LAG 1	10%	decision	1	1	0	1
Augmented Dickey-Fuller TS LAG 1	10%	p-value	0.00	0.00	0.38	0.00
Augmented Dickey-Fuller TS LAG 1	10%	statistic	-15.39	-16.37	-2.43	-16.85
Augmented Dickey-Fuller TS LAG 1	10%	cvalue	-3.13	-3.13	-3.13	-3.13
Augmented Dickey-Fuller TS LAG 2	10%	decision	1.00	1.00	0.00	1.00
Augmented Dickey-Fuller TS LAG 2	10%	p-value	0.00	0.00	0.43	0.00
Augmented Dickey-Fuller TS LAG 2	10%	statistic	-13.09	-12.84	-2.33	-12.47
Augmented Dickey-Fuller TS LAG 2	10%	cvalue	-3.13	-3.13	-3.13	-3.13
Augmented Dickey-Fuller TS LAG 0	5%	decision	1	1	0	1
Augmented Dickey-Fuller TS LAG 0	5%	p-value	0.00	0.00	0.60	0.00
Augmented Dickey-Fuller TS LAG 0	5%	statistic	-18.29	-22.08	-1.98	-16.93
Augmented Dickey-Fuller TS LAG 0	5%	cvalue	-3.42	-3.42	-3.42	-3.42
Augmented Dickey-Fuller TS LAG 1	5%	decision	1	1	0	1

Augmented Dickey-Fuller TS LAG 1	5%	p-value	0.00	0.00	0.38	0.00
Augmented Dickey-Fuller TS LAG 1	5%	statistic	-15.39	-16.37	-2.43	-16.85
Augmented Dickey-Fuller TS LAG 1	5%	cvalue	-3.42	-3.42	-3.42	-3.42
Augmented Dickey-Fuller TS LAG 2	5%	decision	1	1	0	1
Augmented Dickey-Fuller TS LAG 2	5%	p-value	0.00	0.00	0.43	0.00
Augmented Dickey-Fuller TS LAG 2	5%	statistic	-13.09	-12.84	-2.33	-12.47
Augmented Dickey-Fuller TS LAG 2	5%	cvalue	-3.42	-3.42	-3.42	-3.42
Augmented Dickey-Fuller TS LAG 0	1%	decision	1	1	0	1
Augmented Dickey-Fuller TS LAG 0	1%	p-value	0.00	0.00	0.60	0.00
Augmented Dickey-Fuller TS LAG 0	1%	statistic	-18.29	-22.08	-1.98	-16.93
Augmented Dickey-Fuller TS LAG 0	1%	cvalue	-3.98	-3.98	-3.98	-3.98
Augmented Dickey-Fuller TS LAG 1	1%	decision	1	1	0	1
Augmented Dickey-Fuller TS LAG 1	1%	p-value	0.00	0.00	0.38	0.00
Augmented Dickey-Fuller TS LAG 1	1%	statistic	-15.39	-16.37	-2.43	-16.85
Augmented Dickey-Fuller TS LAG 1	1%	cvalue	-3.98	-3.98	-3.98	-3.98
Augmented Dickey-Fuller TS LAG 2	1%	decision	1	1	0	1
Augmented Dickey-Fuller TS LAG 2	1%	p-value	0.00	0.00	0.43	0.00
Augmented Dickey-Fuller TS LAG 2	1%	statistic	-13.09	-12.84	-2.33	-12.47
Augmented Dickey-Fuller TS LAG 2	1%	cvalue	-3.98	-3.98	-3.98	-3.98
Augmented Dickey-Fuller AR LAG 0	10%	decision	1	1	0	1
Augmented Dickey-Fuller AR LAG 0	10%	p-value	0.00	0.00	0.40	0.00
Augmented Dickey-Fuller AR LAG 0	10%	statistic	-18.16	-21.97	-0.68	-16.80
Augmented Dickey-Fuller AR LAG 0	10%	cvalue	-1.62	-1.62	-1.62	-1.62
Augmented Dickey-Fuller AR LAG 1	10%	decision	1	1	0	1
Augmented Dickey-Fuller AR LAG 1	10%	p-value	0.00	0.00	0.37	0.00
Augmented Dickey-Fuller AR LAG 1	10%	statistic	-15.22	-16.23	-0.77	-16.62
Augmented Dickey-Fuller AR LAG 1	10%	cvalue	-1.62	-1.62	-1.62	-1.62
Augmented Dickey-Fuller AR LAG 2	10%	decision	1	1	0	1
Augmented Dickey-Fuller AR LAG 2	10%	p-value	0.00	0.00	0.38	0.00
Augmented Dickey-Fuller AR LAG 2	10%	statistic	-12.90	-12.69	-0.75	-12.25
Augmented Dickey-Fuller AR LAG 2	10%	cvalue	-1.62	-1.62	-1.62	-1.62
Augmented Dickey-Fuller AR LAG 0	5%	decision	1	1	0	1
Augmented Dickey-Fuller AR LAG 0	5%	p-value	0.00	0.00	0.40	0.00
Augmented Dickey-Fuller AR LAG 0	5%	statistic	-18.16	-21.97	-0.68	-16.80
Augmented Dickey-Fuller AR LAG 0	5%	cvalue	-1.94	-1.94	-1.94	-1.94
Augmented Dickey-Fuller AR LAG 1	5%	decision	1	1	0	1
Augmented Dickey-Fuller AR LAG 1	5%	p-value	0.00	0.00	0.37	0.00
Augmented Dickey-Fuller AR LAG 1	5%	statistic	-15.22	-16.23	-0.77	-16.62
Augmented Dickey-Fuller AR LAG 1	5%	cvalue	-1.94	-1.94	-1.94	-1.94
Augmented Dickey-Fuller AR LAG 2	5%	decision	1	1	0	1
Augmented Dickey-Fuller AR LAG 2	5%	p-value	0.00	0.00	0.38	0.00
Augmented Dickey-Fuller AR LAG 2	5%	statistic	-12.90	-12.69	-0.75	-12.25
Augmented Dickey-Fuller AR LAG 2	5%	cvalue	-1.94	-1.94	-1.94	-1.94
Augmented Dickey-Fuller AR LAG 0	1%	decision	1	1	0	1

Augmented Dickey-Fuller AR LAG 0	1%	p-value	0.00	0.00	0.40	0.00
Augmented Dickey-Fuller AR LAG 0	1%	statistic	-18.16	-21.97	-0.68	-16.80
Augmented Dickey-Fuller AR LAG 0	1%	cvalue	-2.57	-2.57	-2.57	-2.57
Augmented Dickey-Fuller AR LAG 1	1%	decision	1	1	0	1
Augmented Dickey-Fuller AR LAG 1	1%	p-value	0.00	0.00	0.37	0.00
Augmented Dickey-Fuller AR LAG 1	1%	statistic	-15.22	-16.23	-0.77	-16.62
Augmented Dickey-Fuller AR LAG 1	1%	cvalue	-2.57	-2.57	-2.57	-2.57
Augmented Dickey-Fuller AR LAG 2	1%	decision	1	1	0	1
Augmented Dickey-Fuller AR LAG 2	1%	p-value	0.00	0.00	0.38	0.00
Augmented Dickey-Fuller AR LAG 2	1%	statistic	-12.90	-12.69	-0.75	-12.25
Augmented Dickey-Fuller AR LAG 2	1%	cvalue	-2.57	-2.57	-2.57	-2.57
Augmented Dickey-Fuller ARD LAG 0	10%	decision	1	1	0	1
Augmented Dickey-Fuller ARD LAG 0	10%	p-value	0.00	0.00	0.70	0.00
Augmented Dickey-Fuller ARD LAG 0	10%	statistic	-18.30	-22.08	-1.10	-16.86
Augmented Dickey-Fuller ARD LAG 0	10%	cvalue	-2.57	-2.57	-2.57	-2.57
Augmented Dickey-Fuller ARD LAG 1	10%	decision	1	1	0	1
Augmented Dickey-Fuller ARD LAG 1	10%	p-value	0.00	0.00	0.50	0.00
Augmented Dickey-Fuller ARD LAG 1	10%	statistic	-15.40	-16.37	-1.55	-16.72
Augmented Dickey-Fuller ARD LAG 1	10%	cvalue	-2.57	-2.57	-2.57	-2.57
Augmented Dickey-Fuller ARD LAG 2	10%	decision	1	1	0	1
Augmented Dickey-Fuller ARD LAG 2	10%	p-value	0.00	0.00	0.54	0.00
Augmented Dickey-Fuller ARD LAG 2	10%	statistic	-13.10	-12.84	-1.45	-12.34
Augmented Dickey-Fuller ARD LAG 2	10%	cvalue	-2.57	-2.57	-2.57	-2.57
Augmented Dickey-Fuller ARD LAG 0	5%	decision	1	1	0	1
Augmented Dickey-Fuller ARD LAG 0	5%	p-value	0.00	0.00	0.70	0.00
Augmented Dickey-Fuller ARD LAG 0	5%	statistic	-18.30	-22.08	-1.10	-16.86
Augmented Dickey-Fuller ARD LAG 0	5%	cvalue	-2.87	-2.87	-2.87	-2.87
Augmented Dickey-Fuller ARD LAG 1	5%	decision	1	1	0	1
Augmented Dickey-Fuller ARD LAG 1	5%	p-value	0.00	0.00	0.50	0.00
Augmented Dickey-Fuller ARD LAG 1	5%	statistic	-15.40	-16.37	-1.55	-16.72
Augmented Dickey-Fuller ARD LAG 1	5%	cvalue	-2.87	-2.87	-2.87	-2.87
Augmented Dickey-Fuller ARD LAG 2	5%	decision	1	1	0	1
Augmented Dickey-Fuller ARD LAG 2	5%	p-value	0.00	0.00	0.54	0.00
Augmented Dickey-Fuller ARD LAG 2	5%	statistic	-13.10	-12.84	-1.45	-12.34
Augmented Dickey-Fuller ARD LAG 2	5%	cvalue	-2.87	-2.87	-2.87	-2.87
Augmented Dickey-Fuller ARD LAG 0	1%	decision	1	1	0	1
Augmented Dickey-Fuller ARD LAG 0	1%	p-value	0.00	0.00	0.70	0.00
Augmented Dickey-Fuller ARD LAG 0	1%	statistic	-18.30	-22.08	-1.10	-16.86
Augmented Dickey-Fuller ARD LAG 0	1%	cvalue	-3.44	-3.44	-3.44	-3.44
Augmented Dickey-Fuller ARD LAG 1	1%	decision	1	1	0	1
Augmented Dickey-Fuller ARD LAG 1	1%	p-value	0.00	0.00	0.50	0.00
Augmented Dickey-Fuller ARD LAG 1	1%	statistic	-15.40	-16.37	-1.55	-16.72
Augmented Dickey-Fuller ARD LAG 1	1%	cvalue	-3.44	-3.44	-3.44	-3.44
Augmented Dickey-Fuller ARD LAG 2	1%	decision	1	1	0	1

Augmented Dickey-Fuller ARD LAG 2	1%	p-value	0.00	0.00	0.54	0.00
Augmented Dickey-Fuller ARD LAG 2	1%	statistic	-13.10	-12.84	-1.45	-12.34
Augmented Dickey-Fuller ARD LAG 2	1%	cvalue	-3.44	-3.44	-3.44	-3.44
Leybourne-McCabe stationarity	10%	decision	0	0	0	0
Leybourne-McCabe stationarity	10%	p-value	0.10	0.10	0.10	0.10
Leybourne-McCabe stationarity	10%	statistic	0.03	0.11	-1654	0.05
Leybourne-McCabe stationarity	10%	cvalue	0.12	0.12	0.12	0.12
Leybourne-McCabe stationarity	5%	decision	0	0	0	0
Leybourne-McCabe stationarity	5%	p-value	0.10	0.10	0.10	0.10
Leybourne-McCabe stationarity	5%	statistic	0.03	0.11	-1654	0.05
Leybourne-McCabe stationarity	5%	cvalue	0.15	0.15	0.15	0.15
Leybourne-McCabe stationarity	1%	decision	0	0	0	0
Leybourne-McCabe stationarity	1%	p-value	0.10	0.10	0.10	0.10
Leybourne-McCabe stationarity	1%	statistic	0.03	0.11	-1654	0.05
Leybourne-McCabe stationarity	1%	cvalue	0.22	0.22	0.22	0.22
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	10%	decision	0	0	1	0
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	10%	p-value	0.10	0.10	0.01	0.10
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	10%	statistic	0.04	0.10	0.36	0.07
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	10%	cvalue	0.12	0.12	0.12	0.12
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	5%	decision	0	0	1	0
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	5%	p-value	0.10	0.10	0.01	0.10
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	5%	statistic	0.04	0.10	0.36	0.07
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	5%	cvalue	0.15	0.15	0.15	0.15
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	1%	decision	0	0	1	0
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	1%	p-value	0.10	0.10	0.01	0.10
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	1%	statistic	0.04	0.10	0.36	0.07
Kwiatkowski, Phillips, Schmidt, and Shin (KPSS)	1%	cvalue	0.22	0.22	0.22	0.22
Ljung-Box Q- residual autocorrelation LAG 5	10%	decision	1	0	1	1
Ljung-Box Q- residual autocorrelation LAG 5	10%	p-value	0.00	0.20	0.00	0.00
Ljung-Box Q- residual autocorrelation LAG 5	10%	statistic	30.02	7.31	2581	56.35
Ljung-Box Q- residual autocorrelation LAG 5	10%	cvalue	9.24	9.24	9.24	9.24

Ljung-Box Q- residual autocorrelation LAG 10	10%	decision	1	0	1	1
Ljung-Box Q- residual autocorrelation LAG 10	10%	p-value	0.00	0.53	0.00	0.00
Ljung-Box Q- residual autocorrelation LAG 10	10%	statistic	33.83	8.99	4914	66.37
Ljung-Box Q- residual autocorrelation LAG 10	10%	cvalue	15.99	15.99	15.99	15.99
Ljung-Box Q- residual autocorrelation LAG 15	10%	decision	1	0	1	1
Ljung-Box Q- residual autocorrelation LAG 15	10%	p-value	0.00	0.58	0.00	0.00
Ljung-Box Q- residual autocorrelation LAG 15	10%	statistic	41.56	13.24	7062	80.19
Ljung-Box Q- residual autocorrelation LAG 15	10%	cvalue	22.31	22.31	22.31	22.31
Ljung-Box Q- residual autocorrelation LAG 5	5%	decision	1	0	1	1
Ljung-Box Q- residual autocorrelation LAG 5	5%	p-value	0.00	0.20	0.00	0.00
Ljung-Box Q- residual autocorrelation LAG 5	5%	statistic	30.02	7.31	2581	56.35
Ljung-Box Q- residual autocorrelation LAG 5	5%	cvalue	11.07	11.07	11.07	11.07
Ljung-Box Q- residual autocorrelation LAG 10	5%	decision	1	0	1	1
Ljung-Box Q- residual autocorrelation LAG 10	5%	p-value	0.00	0.53	0.00	0.00
Ljung-Box Q- residual autocorrelation LAG 10	5%	statistic	33.83	8.99	4914	66.37
Ljung-Box Q- residual autocorrelation LAG 10	5%	cvalue	18.31	18.31	18.31	18.31
Ljung-Box Q- residual autocorrelation LAG 15	5%	decision	1	0	1	1
Ljung-Box Q- residual autocorrelation LAG 15	5%	p-value	0.00	0.58	0.00	0.00
Ljung-Box Q- residual autocorrelation LAG 15	5%	statistic	41.56	13.24	7062	80.19
Ljung-Box Q- residual autocorrelation LAG 15	5%	cvalue	25.00	25.00	25.00	25.00
Ljung-Box Q- residual autocorrelation LAG 5	1%	decision	1	0	1	1
Ljung-Box Q- residual autocorrelation LAG 5	1%	p-value	0.00	0.20	0.00	0.00
Ljung-Box Q- residual autocorrelation LAG 5	1%	statistic	30.02	7.31	2581	56.35
Ljung-Box Q- residual autocorrelation LAG 5	1%	cvalue	15.09	15.09	15.09	15.09

Ljung-Box Q- residual autocorrelation LAG 10	1%	decision	1	0	1	1
Ljung-Box Q- residual autocorrelation LAG 10	1%	p-value	0.00	0.53	0.00	0.00
Ljung-Box Q- residual autocorrelation LAG 10	1%	statistic	33.83	8.99	4914	66.37
Ljung-Box Q- residual autocorrelation LAG 10	1%	cvalue	23.21	23.21	23.21	23.21
Ljung-Box Q- residual autocorrelation LAG 15	1%	decision	1	0	1	1
Ljung-Box Q- residual autocorrelation LAG 15	1%	p-value	0.00	0.58	0.00	0.00
Ljung-Box Q- residual autocorrelation LAG 15	1%	statistic	41.56	13.24	7062	80.19
Ljung-Box Q- residual autocorrelation LAG 15	1%	cvalue	30.58	30.58	30.58	30.58
Phillips-Perron one unit root against TS LAG 4	10%	decision	1	1	0	1
Phillips-Perron one unit root against TS LAG 4	10%	p-value	0.00	0.00	0.45	0.00
Phillips-Perron one unit root against TS LAG 4	10%	statistic	-18.12	-22.08	-2.29	-16.58
Phillips-Perron one unit root against TS LAG 4	10%	cvalue	-3.13	-3.13	-3.13	-3.13
Phillips-Perron one unit root against TS LAG 5	10%	decision	1	1	0	1
Phillips-Perron one unit root against TS LAG 5	10%	p-value	0.00	0.00	0.42	0.00
Phillips-Perron one unit root against TS LAG 5	10%	statistic	-18.09	-22.11	-2.34	-16.56
Phillips-Perron one unit root against TS LAG 5	10%	cvalue	-3.13	-3.13	-3.13	-3.13
Phillips-Perron one unit root against TS LAG 6	10%	decision	1	1	0	1
Phillips-Perron one unit root against TS LAG 6	10%	p-value	0.00	0.00	0.41	0.00
Phillips-Perron one unit root against TS LAG 6	10%	statistic	-18.06	-22.12	-2.37	-16.52
Phillips-Perron one unit root against TS LAG 6	10%	cvalue	-3.13	-3.13	-3.13	-3.13
Phillips-Perron one unit root against TS LAG 4	5%	decision	1	1	0	1
Phillips-Perron one unit root against TS LAG 4	5%	p-value	0.00	0.00	0.45	0.00
Phillips-Perron one unit root against TS LAG 4	5%	statistic	-18.12	-22.08	-2.29	-16.58
Phillips-Perron one unit root against TS LAG 4	5%	cvalue	-3.42	-3.42	-3.42	-3.42

Phillips-Perron one unit root against TS LAG 5	5%	decision	1	1	0	1
Phillips-Perron one unit root against TS LAG 5	5%	p-value	0.00	0.00	0.42	0.00
Phillips-Perron one unit root against TS LAG 5	5%	statistic	-18.09	-22.11	-2.34	-16.56
Phillips-Perron one unit root against TS LAG 5	5%	cvalue	-3.42	-3.42	-3.42	-3.42
Phillips-Perron one unit root against TS LAG 6	5%	decision	1	1	0	1
Phillips-Perron one unit root against TS LAG 6	5%	p-value	0.00	0.00	0.41	0.00
Phillips-Perron one unit root against TS LAG 6	5%	statistic	-18.06	-22.12	-2.37	-16.52
Phillips-Perron one unit root against TS LAG 6	5%	cvalue	-3.42	-3.42	-3.42	-3.42
Phillips-Perron one unit root against TS LAG 4	1%	decision	1	1	0	1
Phillips-Perron one unit root against TS LAG 4	1%	p-value	0.00	0.00	0.45	0.00
Phillips-Perron one unit root against TS LAG 4	1%	statistic	-18.12	-22.08	-2.29	-16.58
Phillips-Perron one unit root against TS LAG 4	1%	cvalue	-3.98	-3.98	-3.98	-3.98
Phillips-Perron one unit root against TS LAG 5	1%	decision	1	1	0	1
Phillips-Perron one unit root against TS LAG 5	1%	p-value	0.00	0.00	0.42	0.00
Phillips-Perron one unit root against TS LAG 5	1%	statistic	-18.09	-22.11	-2.34	-16.56
Phillips-Perron one unit root against TS LAG 5	1%	cvalue	-3.98	-3.98	-3.98	-3.98
Phillips-Perron one unit root against TS LAG 6	1%	decision	1	1	0	1
Phillips-Perron one unit root against TS LAG 6	1%	p-value	0.00	0.00	0.41	0.00
Phillips-Perron one unit root against TS LAG 6	1%	statistic	-18.06	-22.12	-2.37	-16.52
Phillips-Perron one unit root against TS LAG 6	1%	cvalue	-3.98	-3.98	-3.98	-3.98
Phillips-Perron one unit root against AR LAG 4	10%	decision	1	1	0	1
Phillips-Perron one unit root against AR LAG 4	10%	p-value	0.00	0.00	0.38	0.00
Phillips-Perron one unit root against AR LAG 4	10%	statistic	-18.02	-21.98	-0.74	-16.47
Phillips-Perron one unit root against AR LAG 4	10%	cvalue	-1.62	-1.62	-1.62	-1.62

Phillips-Perron one unit root against AR LAG 5	10%	decision	1	1	0	1
Phillips-Perron one unit root against AR LAG 5	10%	p-value	0.00	0.00	0.38	0.00
Phillips-Perron one unit root against AR LAG 5	10%	statistic	-18.00	-22.01	-0.75	-16.46
Phillips-Perron one unit root against AR LAG 5	10%	cvalue	-1.62	-1.62	-1.62	-1.62
Phillips-Perron one unit root against AR LAG 6	10%	decision	1	1	0	1
Phillips-Perron one unit root against AR LAG 6	10%	p-value	0.00	0.00	0.37	0.00
Phillips-Perron one unit root against AR LAG 6	10%	statistic	-17.98	-22.03	-0.76	-16.44
Phillips-Perron one unit root against AR LAG 6	10%	cvalue	-1.62	-1.62	-1.62	-1.62
Phillips-Perron one unit root against AR LAG 4	5%	decision	1	1	0	1
Phillips-Perron one unit root against AR LAG 4	5%	p-value	0.00	0.00	0.38	0.00
Phillips-Perron one unit root against AR LAG 4	5%	statistic	-18.02	-21.98	-0.74	-16.47
Phillips-Perron one unit root against AR LAG 4	5%	cvalue	-1.94	-1.94	-1.94	-1.94
Phillips-Perron one unit root against AR LAG 5	5%	decision	1	1	0	1
Phillips-Perron one unit root against AR LAG 5	5%	p-value	0.00	0.00	0.38	0.00
Phillips-Perron one unit root against AR LAG 5	5%	statistic	-18.00	-22.01	-0.75	-16.46
Phillips-Perron one unit root against AR LAG 5	5%	cvalue	-1.94	-1.94	-1.94	-1.94
Phillips-Perron one unit root against AR LAG 6	5%	decision	1	1	0	1
Phillips-Perron one unit root against AR LAG 6	5%	p-value	0.00	0.00	0.37	0.00
Phillips-Perron one unit root against AR LAG 6	5%	statistic	-17.98	-22.03	-0.76	-16.44
Phillips-Perron one unit root against AR LAG 6	5%	cvalue	-1.94	-1.94	-1.94	-1.94
Phillips-Perron one unit root against AR LAG 4	1%	decision	1	1	0	1
Phillips-Perron one unit root against AR LAG 4	1%	p-value	0.00	0.00	0.38	0.00
Phillips-Perron one unit root against AR LAG 4	1%	statistic	-18.02	-21.98	-0.74	-16.47
Phillips-Perron one unit root against AR LAG 4	1%	cvalue	-2.57	-2.57	-2.57	-2.57

Phillips-Perron one unit root against AR LAG 5	1%	decision	1	1	0	1
Phillips-Perron one unit root against AR LAG 5	1%	p-value	0.00	0.00	0.38	0.00
Phillips-Perron one unit root against AR LAG 5	1%	statistic	-18.00	-22.01	-0.75	-16.46
Phillips-Perron one unit root against AR LAG 5	1%	cvalue	-2.57	-2.57	-2.57	-2.57
Phillips-Perron one unit root against AR LAG 6	1%	decision	1	1	0	1
Phillips-Perron one unit root against AR LAG 6	1%	p-value	0.00	0.00	0.37	0.00
Phillips-Perron one unit root against AR LAG 6	1%	statistic	-17.98	-22.03	-0.76	-16.44
Phillips-Perron one unit root against AR LAG 6	1%	cvalue	-2.57	-2.57	-2.57	-2.57
Phillips-Perron one unit root against ARD LAG 4	10%	decision	1	1	0	1
Phillips-Perron one unit root against ARD LAG 4	10%	p-value	0.00	0.00	0.55	0.00
Phillips-Perron one unit root against ARD LAG 4	10%	statistic	-18.14	-22.09	-1.43	-16.52
Phillips-Perron one unit root against ARD LAG 4	10%	cvalue	-2.57	-2.57	-2.57	-2.57
Phillips-Perron one unit root against ARD LAG 5	10%	decision	1	1	0	1
Phillips-Perron one unit root against ARD LAG 5	10%	p-value	0.00	0.00	0.53	0.00
Phillips-Perron one unit root against ARD LAG 5	10%	statistic	-18.11	-22.12	-1.48	-16.51
Phillips-Perron one unit root against ARD LAG 5	10%	cvalue	-2.57	-2.57	-2.57	-2.57
Phillips-Perron one unit root against ARD LAG 6	10%	decision	1	1	0	1
Phillips-Perron one unit root against ARD LAG 6	10%	p-value	0.00	0.00	0.52	0.00
Phillips-Perron one unit root against ARD LAG 6	10%	statistic	-18.08	-22.13	-1.50	-16.47
Phillips-Perron one unit root against ARD LAG 6	10%	cvalue	-2.57	-2.57	-2.57	-2.57
Phillips-Perron one unit root against ARD LAG 4	5%	decision	1	1	0	1
Phillips-Perron one unit root against ARD LAG 4	5%	p-value	0.00	0.00	0.55	0.00
Phillips-Perron one unit root against ARD LAG 4	5%	statistic	-18.14	-22.09	-1.43	-16.52
Phillips-Perron one unit root against ARD LAG 4	5%	cvalue	-2.87	-2.87	-2.87	-2.87

Phillips-Perron one unit root against ARD LAG 5	5%	decision	1	1	0	1
Phillips-Perron one unit root against ARD LAG 5	5%	p-value	0.00	0.00	0.53	0.00
Phillips-Perron one unit root against ARD LAG 5	5%	statistic	-18.11	-22.12	-1.48	-16.51
Phillips-Perron one unit root against ARD LAG 5	5%	cvalue	-2.87	-2.87	-2.87	-2.87
Phillips-Perron one unit root against ARD LAG 6	5%	decision	1	1	0	1
Phillips-Perron one unit root against ARD LAG 6	5%	p-value	0.00	0.00	0.52	0.00
Phillips-Perron one unit root against ARD LAG 6	5%	statistic	-18.08	-22.13	-1.50	-16.47
Phillips-Perron one unit root against ARD LAG 6	5%	cvalue	-2.87	-2.87	-2.87	-2.87
Phillips-Perron one unit root against ARD LAG 4	1%	decision	1	1	0	1
Phillips-Perron one unit root against ARD LAG 4	1%	p-value	0.00	0.00	0.55	0.00
Phillips-Perron one unit root against ARD LAG 4	1%	statistic	-18.14	-22.09	-1.43	-16.52
Phillips-Perron one unit root against ARD LAG 4	1%	cvalue	-3.44	-3.44	-3.44	-3.44
Phillips-Perron one unit root against ARD LAG 5	1%	decision	1	1	0	1
Phillips-Perron one unit root against ARD LAG 5	1%	p-value	0.00	0.00	0.53	0.00
Phillips-Perron one unit root against ARD LAG 5	1%	statistic	-18.11	-22.12	-1.48	-16.51
Phillips-Perron one unit root against ARD LAG 5	1%	cvalue	-3.44	-3.44	-3.44	-3.44
Phillips-Perron one unit root against ARD LAG 6	1%	decision	1	1	0	1
Phillips-Perron one unit root against ARD LAG 6	1%	p-value	0.00	0.00	0.52	0.00
Phillips-Perron one unit root against ARD LAG 6	1%	statistic	-18.08	-22.13	-1.50	-16.47
Phillips-Perron one unit root against ARD LAG 6	1%	cvalue	-3.44	-3.44	-3.44	-3.44
Variance ratio random walk	10%	decision	1	1	1	1
Variance ratio random walk	10%	p-value	0.00	0.00	0.00	0.00
Variance ratio random walk	10%	statistic	-5.38	-7.50	4.32	-3.14
Variance ratio random walk	10%	cvalue	1.64	1.64	1.64	1.64
Variance ratio random walk	5%	decision	1	1	1	1
Variance ratio random walk	5%	p-value	0.00	0.00	0.00	0.00
Variance ratio random walk	5%	statistic	-5.38	-7.50	4.32	-3.14

Variance ratio random walk	5%	cvalue	1.96	1.96	1.96	1.96
Variance ratio random walk	1%	decision	1	1	1	1
Variance ratio random walk	1%	p-value	0.00	0.00	0.00	0.00
Variance ratio random walk	1%	statistic	-5.38	-7.50	4.32	-3.14
Variance ratio random walk	1%	cvalue	2.58	2.58	2.58	2.58

Ten different stationarity tests have been performed with combinations of different lags and significance levels in the attempt to check if each one of the three asset class time series of returns would be stationary.

The Augmented Dickey-Fuller TS tests the null hypothesis of a unit root against the trend-stationary alternative, the Augmented Dickey-Fuller AR test the null hypothesis of a unit root against the autoregressive alternative while the Augmented Dickey-Fuller ARD tests the null hypothesis of a unit root against the autoregressive with drift alternative. For all the three different versions of the Augmented Dickey-Fuller test have been performed three tests at lags 0,1 and 2 and significance levels of 10%,5% and 1%. The null hypothesis of a unit root has been rejected for the four time series in all the tests except for the div_y time series which appeared to be not stationary. The Leybourne-McCabe stationarity test assesses the null hypothesis that a univariate time series is a trend stationary AR(p) process, against the alternative that it is a nonstationary ARIMA(p,1,1) process, where p represents the autoregressive order, in this case I chose p equal to 1. The test has been performed at significance levels of 10%,5% and 1%. The null hypothesis failed to be rejected for all the time series at any significance level. The Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) assesses the null hypothesis that a univariate time series is trend stationary against the alternative that it is a nonstationary unit root process. The authors of the test suggest that a number of lags on the order of square root of T, where T is the sample size, is often satisfactory under both the null and the alternative. Given that the length of the four time series analyzed is 540 I chose the number of lags equal to 23. The test has been performed at significance levels of 10%,5% and 1%. The null hypothesis of a unit root failed to be rejected in all the tests except for the div_y time series which appeared to be not stationary. The Phillips-Perron TS tests null hypothesis of a unit root against the trend-stationary alternative, the Phillips-Perron AR tests

null hypothesis of a unit root against the autoregressive alternative while the Phillips-Perron ARD tests null hypothesis of a unit root against the autoregressive with drift alternative. The Phillips and Perron's test statistics can be viewed as Dickey–Fuller statistics that have been made robust to serial correlation by using the Newey–West (1987) heteroskedasticity and autocorrelation-consistent covariance matrix estimator. The lags, which is an input of the statistics test, represent the number of autocovariance lags to include in the Newey-West estimator of the long-run variance. The Newey-West estimator (1987)[5] is consistent if the number of lags are $O(T^{1/4})$, where T is the effective sample size; in the case of this study the numbers of lags have been chosen to be 4,5 and 6. The test has been performed at significance levels of 10%,5% and 1%. The null hypothesis of a unit root has been rejected for the four time series in all the tests except for the div_y time series which appeared to be not stationary. The variance ratio test assesses the null hypothesis that a univariate time series is a random walk. The test has been performed at significance levels of 10%,5% and 1% and then null hypothesis has been rejected at any significance level for all the four time series.

In the following passage the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) are plotted for the four time series. Autocorrelation is the linear dependence of a variable with itself at two points in time, it measures the correlation between yt and yt + k, where k = 0,...,K and yt is a stochastic process . Correlation between two variables can result from a mutual linear dependence on other variables. Partial autocorrelation is the autocorrelation between yt and yt–h after removing any linear dependence on y1, y2, ..., yt–h+1. As suggested by Box, Jenkins, and Reinsel (1994)[6] the correlation at each lag is scaled by the sample variance so that the autocorrelation and the partial autocorrelation at lag 0 is unity.

[5] Whitney K. Newey and Kenneth D. West in their paper "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix" (1987)

[6] Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. "Time Series Analysis: Forecasting and Control. 3rd ed. Englewood Cliffs, NJ: Prentice Hall" (1994)

Figure 32 lo20 sample autocorrelation and sample partial autocorrelation

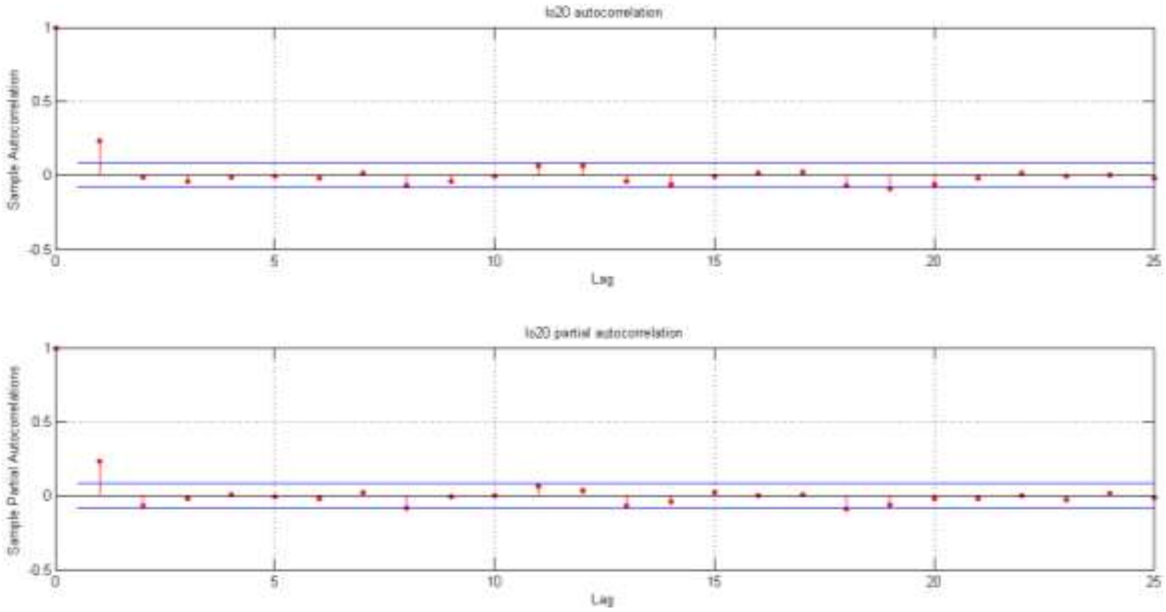


Figure 33 hi20 sample autocorrelation and sample partial autocorrelation

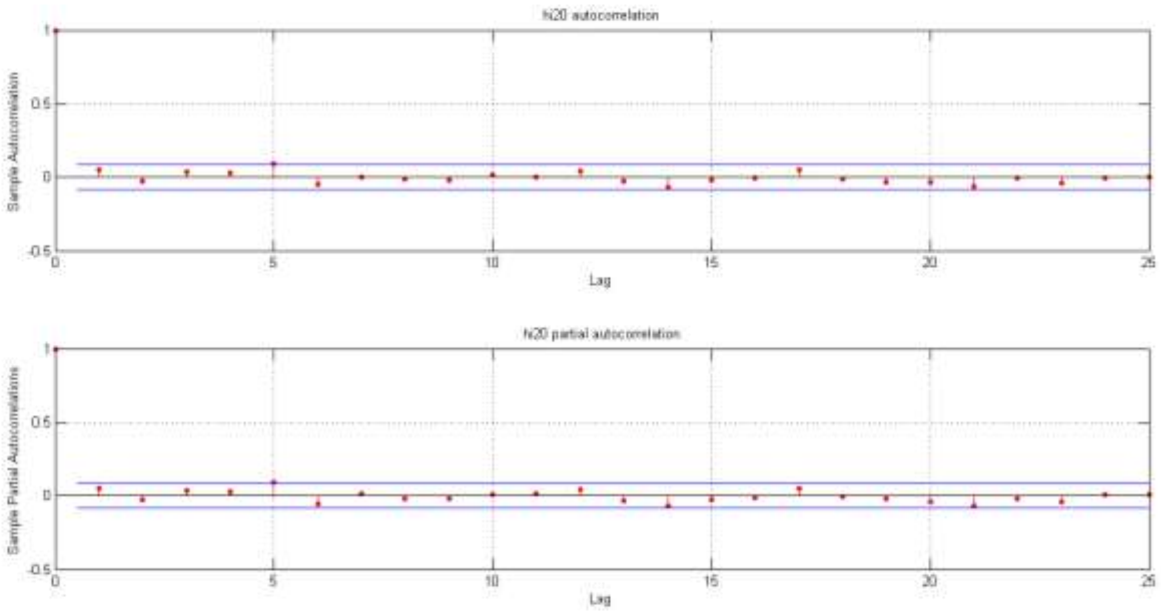


Figure 34 tbond sample autocorrelation and sample partial autocorrelation

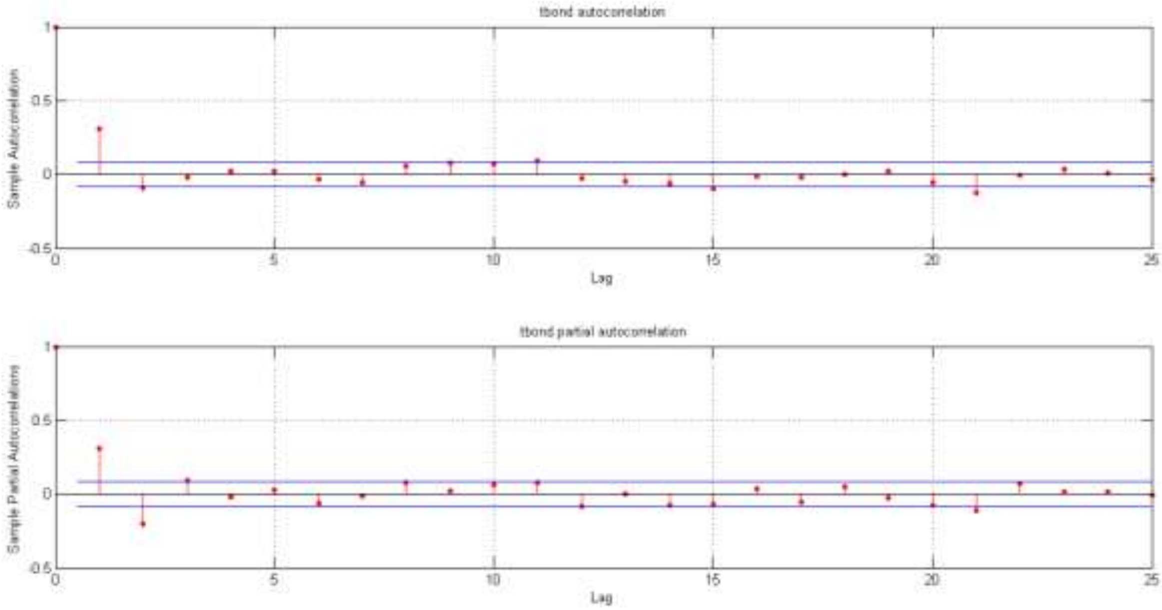
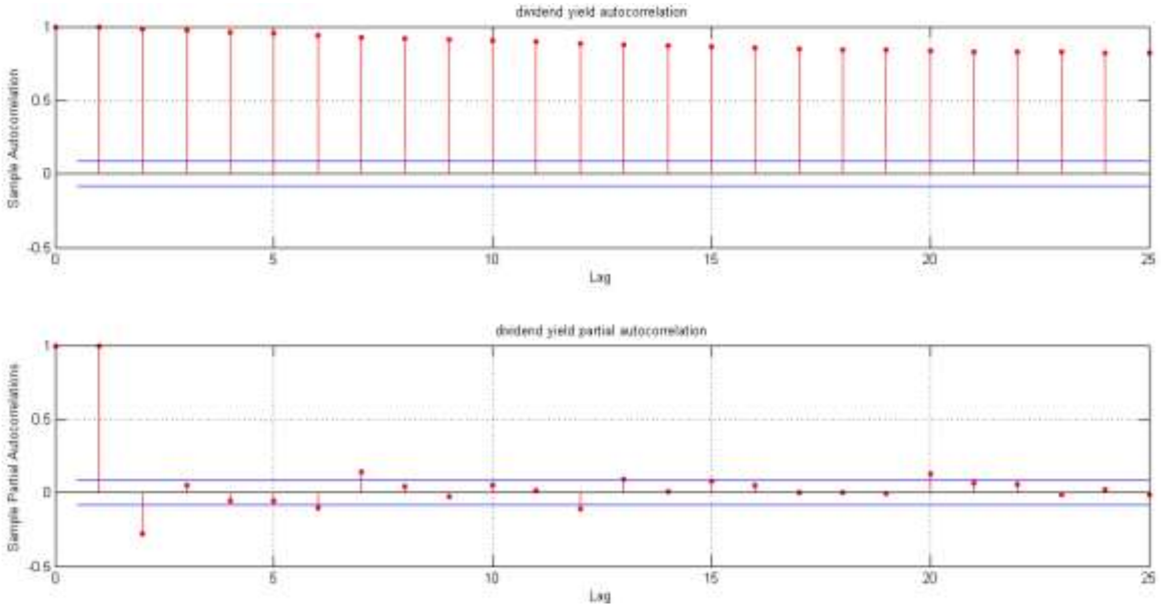


Figure 35 div_y sample autocorrelation and sample partial autocorrelation



The Ljung-Box Q-test is a quantitative way to test for autocorrelation at multiple lags jointly. The test assesses whether a series of residuals exhibits no autocorrelation for a fixed number of lags, against the alternative that some autocorrelation coefficient is nonzero. The null hypothesis for this test is that the first m autocorrelations are jointly zero. Under the null hypothesis, the asymptotic distribution of the test statistics is chi-square with degrees of freedom equal to the sample size. I performed the test at lags 15,10 and 5 and significance levels of

10%,5% and 1%. The null hypothesis failed to be rejected only for hi20 at any lags value and significance level.

From the examination of the results presented in this paragraph emerges that the three asset classes time series are stationary and may need to be model by an autoregressive model while the div_y time series is not stationary.

CHAPTER 3 – MODEL ESTIMATION

3.1 The basic idea behind regime switching models

The basic idea is simple, financial operators care for detecting and forecasting instability in statistical relationships. A naïve approach is to model such instability using dummy variables in “regression-type” analysis, in other words it can be said that a regime applies before the break and the other after the break. Unfortunately this approach is feasible only ex-post because the current regime and its duration are not observable but only forecastable. Econometricians have developed methods in which instability is stochastic, it has a structure, and it can be predicted. Suppose that the historical behavior of returns on some asset can be described with a first-order autoregression process AR(1): the model can be adequate until a certain date when its forecast accuracy declined, at this point the idea would be to simply estimate again the parameter of the model if this break can be observed, otherwise if the break is not observable the operators face uncertainty and a possible model misspecification which can lead to a poor forecast accuracy. In this framework the change that occurred at a certain date was considered “deterministic” by operators and anyone would have been able to predict with certainty looking ahead and anyone would have estimated the parameter of the model again just after the break occurrence. Instead in reality there must have been some imperfectly predictable forces that produced the change. Hence, the econometricians developed a more elegant approach that assumes there is some larger model encompassing both regimes. If our model is still the AR(1) now it would have switching parameters whose values are conditional on the regimes the process is at certain time. This means that the parameters of the AR(1) are now time-variant parameters which depend on the value of the current regime/state of the process. A complete description of the probability law governing the observed data would then require a probabilistic model of what caused the change from one regime to another. More generally, to specify and estimate a regime switching model will mean to specify a structural model for the temporal dependence of the regimes state variable.

At this point, the problem becomes to choose what kind of stochastic process can be specified for the state variable; three mainly different solutions have been

proposed: threshold models in which the state variable can assume k distinct values in dependence of the values as of time t of some threshold variable x, smooth transition models in which the state variable comes from some discrete probability distribution that can take k distinct values and the probability depends on the values as of time t of some threshold variable x, it is called smooth because the variable x no longer determistically determines the state, but simply the cumulative distribution function of the regime; Markov switching models in which the state variable is unobservable from a discrete, first-order, k-state, irreducible, ergodic Markov chain. Markov switching models are of intermediate generality between threshold and smooth transition models. Discrete means that the state variable may only take k distinct values. First-order Markov chain means that the probability of the realization of a certain value *i* of the state variable at time *t* only depends on the value of the state variable at time *t-1*. Irreducible means that there is no absorbing state, i.e. no special regime *i* such that $\mathrm{Pr}(\mathrm{S_t} = i \,|\, \mathrm{S_{t-1}} = \mathrm{j}) = 0$ for all $i = 1,2,\ldots,k$ in other words no regimes exists such that the chain gets trapped into it. Ergodic means that the chain has a long-run mean; in fact, defining δ_t as a *(k x 1)* vector made of zeros except the *j-th* element that equals 1 to signal $\mathrm{S_t} = j$ and 0 otherwise, we have: $\delta_{t+1} = \mathrm{P}'\delta_t + \mathrm{v}_{t+1}$ where P is the transition matrix that collects the quantities:

(1) $\mathrm{Pr}(\mathrm{S_t} = i \,|\, \mathrm{S_{t-1}} = j) = \mathrm{p}_{ij}$,

and v_{t+1} is some error term. Notice that such representation implied that the forecast is: $\mathrm{E}[\delta_{t+\mathrm{T}}|\delta_t, \delta_{t-1}, \delta_{t-2}, \ldots] = (\mathrm{P}')^{\mathrm{T}}\delta_t$. A Markov chain is ergodic if and only if: $\mathrm{plim}\,\mathrm{E}[\delta_{t+\mathrm{T}}|\delta_t, \delta_{t-1}, \delta_{t-2}, \ldots] = \mathrm{plim}\,(\mathrm{P}')^{\mathrm{T}}\delta_t = \pi$ where plim means "limit probability" as $\mathrm{T} \to \infty$ and π is called vector of unconditional or ergodic probabilities (also called ξ_∞). One useful implication from all of this is that is possible to forecast the probability of a regime T steps ahead using the following formula $(\mathrm{P}')^{\mathrm{T}}\delta_t$ because $\mathrm{Pr}(\mathrm{S_{t+T}} | \,\mathrm{S_t}) = \mathrm{E}[\delta_{t+\mathrm{T}}|\delta_t]$.

3.2 Filtered and smoothed regimes probabilities

The key feature of Markov switching models is that they allow you to endogenously make inferences on the regimes based on the data, for instance, in

the case with two regimes/states, because the regimes are unobservable (also said, latent) the econometrician observes the returns directly but can only make an inference about the value of the state variable S_t based on what he sees happening with the returns. With two regimes this inferences will take the form of two probabilities:

(2) $\xi_{1t} \equiv \Pr(S_t = 1|\Im_t, \theta)$ and $\xi_{2t} \equiv \Pr(S_t = 2|\Im_t, \theta)$

where $\Im_t$ is all the past information up on time *t* and θ is the vector collecting the parameters to be estimated. ξ_{1t} and ξ_{2t} are called "filtered state (regime) probabilities" because they depend on "real time" information $\Im_t$ and $\xi_{1t} + \xi_{2t} = 1$. A filtered probability is the best inference of the current state, based on real time information. The Hamilton's filter is used to calculate the filtered probabilities of each state based on the arrival of new information.

Here is illustrated how to compute filtered probabilities. The following formulas regard the univariate case but are easily extended to deal with the multivariate case.

Assume that ξ_{1t-1} is already known and that $f(R_t|S_t = j, \Im_{t-1}; \theta)$ is the log-likelihood of the time *t* observation R_t conditional on state $S_t = j$. Then the conditional density of the time *t* observation is:

(3)
$$f(R_t|\Im_{t-1}; \theta) = \Pr(R_t|\Im_{t-1}; \theta) = \xi_{1t-1} p_{11} f(R_t|S_t = 1, \Im_{t-1}; \theta) + \xi_{1t-1} p_{21} f(R_t|S_t = 2, \Im_{t-1}; \theta) + (1 - \xi_{1t-1}) p_{12} f(R_t|S_t = 1, \Im_{t-1}; \theta) + (1 - \xi_{1t-1}) p_{22} f(R_t|S_t = 2, \Im_{t-1}; \theta)$$

At this point, by an application of Bayes' rule, can be obtain:

(4)
$$\xi_{jt} = \frac{\xi_{1t-1} p_{j1} f(R_t|S_t = j, \Im_{t-1};\ \theta) + \xi_{2t-1} p_{j2} f(R_t|S_t = j, \Im_{t-1};\ \theta)}{f(R_t|\Im_{t-1};\ \theta)} \quad for\ j = 1,2$$

In other words the probability of being in regime j at time *t* is the ratio between the probability of reaching j from $S_{t-1} = 1$ plus the probability of reaching j from $S_{t-1} =$

2 and the total probability of R_t given past information. There are various techniques to initialize ξ_{10} in order to compute ξ_{11} and ξ_{12} : if the Markov chain is presumed to be ergodic one can use the unconditional probabilities:

(5) $$\xi_{1\infty} = \xi_{10} = \frac{1-p_{22}}{2-p_{11}-p_{22}} \quad \xi_{2\infty} = \xi_{20} = \frac{1-p_{11}}{2-p_{11}-p_{22}}$$

Other common choices are $\xi_{10} = \xi_{20} = 0.5$ or ξ_{10} and therefore $\xi_{20} = (1 - \xi_{10})$ can be estimated by MLE as if they were themselves parameters.

One is also often interested in forming an inference about what regime the economy was in at date *t* based on observation obtained through a later date *T*: $\xi_{1t|T} \equiv \Pr(S_t = 1|\Im_T; \theta)$ and $\xi_{2t|T} \equiv \Pr(S_t = 2|\Im_T; \theta)$. These are referred to as "smoothed probabilities". The obvious difference between smoothed and filtered probabilities is that the former ones use all the information in the sample and as such they represent ex-post measures. A filtered probability provides instead a recursive, real time assessment on the current state.
The average duration of state *i*, in a two states model, is equal to:

(6) $$\frac{1}{1-p_{ii}}$$

3.3 The estimation process of a Markov regime switching model

Different econometric methods can be used to estimate regime switching models: maximum-likelihood and EM algorithms are outlined by Hamilton (1988; 1990). The maximum likelihood algorithm involves a Bayesian updating procedure, which infers the probability of being in a regime given all available information up until that time. An alternative to maximum likelihood estimation is Gibbs sampling, which was developed for regime switching models by Albert and Chib (1993) and Kim and Nelson (1999), in this approach both the parameters and the Markov switching are treated as random variables.

An important issue in estimating regime switching models is specifying the number of regimes. This is often difficult to determine from data, and as far as possible the choice should be based on economic arguments. Such decision can be difficult given that the regimes themselves are often thought of as approximation to underlying states that are, unobserved. It is not uncommon to simply fix the

number of regime at some values, typically two (bear and bull markets), rather than basing the decision on econometric tests. The reason is that tests for the number of regimes are typically difficult to implement because they do not follow standard distributions. Under the null of a single regime, the parameters of the other regime are not identified, so there are unidentified nuisance parameters. This means that conventional likelihood ratio tests are not asymptotically Chi-squared distributed.

The algorithm for the estimation of Markov switching model proposed by Kim and Nelson, instead of proceeding forward in a recursive fashion, proceeds backwards, starting from the fact that for *T* = *t*, filtered and smoothed probabilities will be identical. The estimation of Markov switching models performed by MLE possesses optimal properties. This means that under standard regularity conditions, the ML estimator of θ will be: consistent, asymptotically normal, the most efficient in a wide class of estimators. The conditional log-likelihood of the observed data arises rather naturally from the recursive calculation of filtered probabilities

(7) $\ln f(R_1, R_2, \dots, R_{\mathrm{T}} | R_0; \boldsymbol{\theta}) = \sum_{\mathrm{t}=1}^{\mathrm{T}} \ln f(R_{\mathrm{t}} | \Im_{\mathrm{t}-1}; \boldsymbol{\theta})$.

The estimation procedure by maximum-likelihood is here described assuming the existence of two regimes. We observe R_t directly but we can only make an inference about the value of $\mathrm{S_t}$ based on what we see happening with R_t.The inference will take the form of the two probabilities $\xi_{\mathrm{jt}} = \Pr(S_{\mathrm{t}} = j | \Im_{\mathrm{t}}, \boldsymbol{\theta})$ for $j = 1{,}2$ where $\Im_{\mathrm{t}}$ is the set of observations obtained as of date *t* and θ represents the vector of population parameters. The inference is performed iteratively for $t = 1, 2, \dots, T$ with step *t* accepting as input the values $\xi_{\mathrm{it}-1} = \Pr(\mathrm{S_{t-1}} = i | \Im_{\mathrm{t}-1}, \boldsymbol{\theta})$ for $i = 1{,}2$ and producing as output $\xi_{1\mathrm{t}}$ and $\xi_{2\mathrm{t}}$. The densities under the two regimes are denoted as

(8) $\eta_{jt} = f(R_t | \mathrm{S_t} = j, \Im_{\mathrm{t}-1}; \boldsymbol{\theta})$ for $j = 1{,}2$.

Given the input $\xi_{\mathrm{it}-1}$ it is possible to calculate the conditional density of the *t*-th observation from:

(9)[7] $f(R_t|\mathfrak{I}_{\mathrm{t-1}};\boldsymbol{\theta}) = \sum_{j=1}^{2}\sum_{i=1}^{2} \mathrm{p}_{ji}\, \xi_{\mathrm{it-1}}\eta_{jt}$

and the desired output is then:

(10) $\xi_{\mathrm{jt}} = \frac{\sum_{i=1}^{2} \mathrm{p}_{ji}\xi_{it-1}\eta_{jt}}{f(R_t|\mathfrak{I}_{\mathrm{t-1}};\boldsymbol{\theta})}$

Executing this iteration will lead to evaluate the sample conditional log-likelihood of the observed data for the specified value of $\boldsymbol{\theta}$

(7) $\ln f(R_1, R_2, \dots, R_T|R_0;\boldsymbol{\theta}) = \sum_{\mathrm{t}=1}^{T} \ln f(R_t|\mathfrak{I}_{\mathrm{t-1}};\boldsymbol{\theta})$

An estimation of $\boldsymbol{\theta}$ can then be obtained by maximizing the log likelihood of the observed data by numerical optimization. Several options are available for the value ξ_{i0} to use to start these iterative process, some of them have been already described in this paragraph.

The previous estimation steps can be easily generalized to consider a vector of observation $\boldsymbol{R_t}$ at time *t* and then a multivariate model. Let $\boldsymbol{\mathfrak{I}_{t-1}} = \{\boldsymbol{R_t}, \boldsymbol{R_{t-1}}, \dots, \boldsymbol{R_1}\}$ be the observation through date *t*, $\mathbf{P}$ be an *(n x n)* matrix whose row *j*, column *i* element is the transition probability p_{ij}, $\boldsymbol{\eta_t}$ be an *(n x 1)* vector whose *j*-th element $f(\boldsymbol{R_t}|\mathrm{S_t} = j, \mathfrak{I}_{\mathrm{t-1}};\boldsymbol{\theta})$ is the density in regime *j*, and $\hat{\boldsymbol{\xi}}_{\mathbf{t|t}}$ an *(n x 1)* vector whose *j*-th element is $\Pr(\mathrm{S_t} = j|\mathfrak{I}_{\mathrm{t}};\boldsymbol{\theta})$. Then the conditional density of the *t*-th observation is:

(11)[8] $f(\boldsymbol{R_t}|\mathfrak{I}_{\mathrm{t-1}};\boldsymbol{\theta}) = \mathbf{1}'\left(\mathbf{P}\hat{\boldsymbol{\xi}}_{\mathbf{t-1|t-1}} \odot \boldsymbol{\eta_t}\right)$

where $\hat{\boldsymbol{\xi}}_{\mathbf{t|t-1}} = \mathbf{P}\hat{\boldsymbol{\xi}}_{\mathbf{t-1|t-1}}$ and desired output is then:

(12) $\hat{\boldsymbol{\xi}}_{\mathbf{t|t}} = \frac{\mathbf{P}\hat{\boldsymbol{\xi}}_{\mathbf{t-1|t-1}} \odot \boldsymbol{\eta_t}}{f(\boldsymbol{R_t}|\mathfrak{I}_{\mathrm{t-1}};\boldsymbol{\theta})}$

where $\mathbf{1}$ is a *(n x 1)* vector all of whose elements are unity and $\odot$ denotes element-by-element multiplication.

[7] James D. Hamilton “Regime-Switching Models” (2005)
[8] James D. Hamilton “Regime-Switching Models. Palgrave Dictionary of Economics” (2005) and Birger Nilsson and Andreas Graflund “Dynamic Portfolio Selection: The Relevance of Switching Regimes and Investment Horizon” (2001)

3.4 Overview of the most common multivariate Markov regime switching models

In this paragraph I am going to briefly describe some of the most common and employed Markov Switching model specifications. The Markov Switching model allows for a great variety of specifications, especially the MSVAR family which nests all the models illustrated in this paragraph. According to the most common notation the presence and the configuration of the regime-dependent parameters is indicated by the presence of particular capital letters in the name of the model specification. The general term MMS(M) stands for Multivariate Markov Switching and the mutable string (M) assumes values in accordance to the configuration of the regime-dependent parameters; the possible (M) values are here explained:

- I Markov switching *intercept* term;
- A Markov switching *autoregressive parameters;*
- H Markov switching *heteroskedasticity*

For example if a multivariate model specification is characterized to accommodate heteroskedasticity through the presence of a regime-dependent covariance matrix and has a regime-dependent intercept term but invariant autoregressive parameters, the model according to the common notation take the name MMSIH-VAR(k,p).

The Markov Switching Vector Autoregressive model MSVAR is a general class of models which nests the standard VAR model but additionally accounts for nonlinear regime shifts. These models are particularly useful when one extends them to capture the dynamics not of the returns only, but also of some vector variables that also collects one or a number of predictors of subsequent asset returns. Many authors adopted this family of Markov switching model to predict stocks and bonds returns using a number of macroeconomic variables as predictors. Many authors have also showed that while a simple VAR(1) model does not produce useful predictions, models with two or more regimes/states manage to be rather useful. MSVAR models are also particularly suitable to model and study contagion dynamics tanks to the presence of the autoregressive

coefficient matrix which might also be time-dependent in certain model configurations.

There are several models configuration nested in the general MSVAR family model, the most relevant are illustrated in the following subparagraph.

3.4.1 MMSIAH(k,p)

The model, according to different notations, is also known as MMSIAH(k)-VAR(p) or MSIVARH(k,p). This model configuration can be considered the most general and complete among all the MSVAR family models; in fact all of its terms – intercept, autoregressive coefficients matrix and covariance matrix – are regime-dependent. It is able to capture and model in a multivariate framework many of the financial time series features. In a MMSIAH(k,p) model there are three types of contagion effects: a simultaneous one through the off-diagonal elements of the variance and covariance matrix that captures the dynamics across regimes of correlations, a dynamic and linear one through the VAR coefficients, a dynamic and nonlinear one through the fact that the regime variables that drives the process of all variables are common to all variables.

When k=1 a MMSIAH(1)-VAR(p) becomes a VAR(p) model with Gaussian shocks. A VAR model is just a multivariate extension of a standard autoregressive model, in which lags of variable i may in principle affects the subsequent value of j. The joint distribution of a vector of *n* returns $\mathrm{r_t} = [\mathrm{r_{1t}}\, \mathrm{r_{2t}} \dots \mathrm{r_{nt}}]'$ can be modeled as a multivariate regime switching process driven by a common discrete state variables, S_t, that takes integer values between 1 and k:

$$(13) \quad \mathbf{r_t} = \boldsymbol{\mu}_{\mathbf{S_t}} + \sum_{j=1}^{\mathrm{p}} \boldsymbol{A}_{\boldsymbol{j},\mathbf{S_t}} \mathbf{r}_{\mathbf{t}-\boldsymbol{j}} + \boldsymbol{\varepsilon}_{\mathbf{t}}$$

where $\boldsymbol{\mu}_{\mathbf{S_t}} = [\mu_{1\mathrm{S_t}} \dots \mu_{\mathrm{nS_t}}]'$ is a vector of mean returns in state $\mathrm{S_t}$, $\boldsymbol{A}_{\boldsymbol{j},\mathbf{S_t}}$ is a (*n x n)* matrix of autoregressive and regression coefficients at lag *j* in state $\mathrm{S_t}$ and $\boldsymbol{\varepsilon}_{\mathbf{t}} = [\varepsilon_{1\mathrm{t}} \dots \varepsilon_{\mathrm{nt}}]' \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{S_t}})$ is the vector of return innovations that are assumed to be joint normally distributed with zero mean and state-specific covariance matrix $\boldsymbol{\Sigma}_{\mathbf{S_t}}$. Innovations to returns are thus drawn from a Gaussian mixture distribution that is known to provide a flexible approximation to a wide class of distributions. The state dependence of the covariance matrix captures the possibility of

heteroskedastic shocks to asset returns and a non diagonal $\boldsymbol{\Sigma}_{\mathbf{S_t}}$ makes the asset returns simultaneously cross-correlated.

Each state is the realization of a first-order Markov chain governed by the *(k x k)* transition probability matrix, P, with generic element p_{ij} defined as $\Pr(\mathrm{S_t} = i \mid \mathrm{S_{t-1}} = j) = \mathrm{p}_{ij} \quad i, j = 1, \dots, k$. In the framework of this model the estimation method allows $\mathrm{S_t}$ to be unobservable and treated as a latent variable.

The version of the model that has been just described above is characterized to have a full covariance matrix estimated from the data, however a less sophisticated version in which the structure of the covariance matrix is diagonal is also possible; in this version the covariance between the residuals of the different equations are not allowed. In the diagonal covariance matrix version of the model there is no evidence of a simultaneous contagious effect one through the off-diagonal elements of the variance and covariance matrix that captures the dynamics across regimes of correlations because these covariance elements are equal to 0. The model can be extended to incorporate an *(l x 1)* vector of predictor variables such as the dividend yield:

$$(14) \quad \boldsymbol{y}_{\mathbf{t}} = \begin{pmatrix} \boldsymbol{\mu}_{\mathbf{S_t}} \\ \boldsymbol{\mu}_{\mathbf{zS_t}} \end{pmatrix} + \sum_{j=1}^{\mathrm{p}} \boldsymbol{A}^{*}_{\boldsymbol{j},\mathbf{S_t}} \mathbf{y}_{\mathbf{t}-\boldsymbol{j}} + \begin{pmatrix} \boldsymbol{\varepsilon}_{\mathbf{t}} \\ \boldsymbol{\varepsilon}_{\mathbf{zt}} \end{pmatrix}$$

the same model can be represented alternatively in the following way:

$$(15) \quad \begin{pmatrix} \mathbf{r_t} \\ \mathbf{z_t} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{\mathbf{S_t}} \\ \boldsymbol{\mu}_{\mathbf{zS_t}} \end{pmatrix} + \sum_{j=1}^{\mathrm{p}} \boldsymbol{A}^{*}_{\boldsymbol{j},\mathbf{S_t}} \begin{pmatrix} \mathbf{r}_{\mathbf{t}-\boldsymbol{j}} \\ \mathbf{z}_{\mathbf{t}-\boldsymbol{j}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_{\mathbf{t}} \\ \boldsymbol{\varepsilon}_{\mathbf{zt}} \end{pmatrix}$$

where $\boldsymbol{y}_{\mathbf{t}} = (\mathbf{r}_{\mathbf{t}}' \ \mathbf{z}_{\mathbf{t}}')'$ is a *(l + n) x 1*, $\boldsymbol{\mu}_{\mathbf{zS_t}} = [\mu_{z1\mathrm{S_t}} \dots \mu_{\mathrm{lS_t}}]'$ is the intercept vector for $\mathbf{z_t}$ in state $\mathrm{S_t}$, $\{\boldsymbol{A}^{*}_{\boldsymbol{j},\mathbf{S_t}}\}_{j=1}^{\mathrm{p}}$ are *(l + n) x (l + n)* matrices of autoregressive and regression coefficients in state $\mathrm{S_t}$ and $[\varepsilon_{\mathrm{t}}' \dots \varepsilon_{\mathrm{zt}}']' \sim N(\mathbf{0}, \boldsymbol{\Sigma}^{*}_{\mathbf{S_t}})$, where $\boldsymbol{\Sigma}^{*}_{\mathbf{S_t}}$ is an *(l + n) x (l + n)* covariance matrix. This model allows for predictability in returns through the lagged values of $\mathbf{z_t}$. The relationship between stock returns and the dividend yield is linear within a given regime but the model is capable of tracking a non-linear relationship between asset returns and the dividend yield since the coefficient on the yield varies across regimes and the regime probabilities changes as well. This family of models even in the absence of autoregressive terms or predictor

variables implies time-varying investment opportunities. For example, the conditional mean of asset returns is an average of the vector of mean returns, $\boldsymbol{\mu}_{\mathrm{S_t}}$, weighted by the filtered state probabilities $[\Pr(\mathrm{S_t}=1|\mathfrak{I}_\mathrm{t}) \ldots \Pr(\mathrm{S_t}=k|\mathfrak{I}_\mathrm{t})]'$, conditional on information available at time t, $\mathfrak{I}_\mathrm{t}$. Since these state probabilities vary over time, the expected return will also change. These same dynamic is applicable to higher order moments of the returns distribution.

The model is stationary if the absence of roots outside the unit circle is verified for the matrices of autoregressive coefficients. Ang and Bekaert (2002) have showed that formally, it is just sufficient for such a condition to be verified in at least one of the k regimes.

Likewise the version without any predictor variables, also the less sophisticated version with predictor variables can be characterized by a diagonal structure of the covariance matrix. In this version the covariance between the residuals of the different equations are not allowed thus there is no evidence of a simultaneous contagious effect through the off-diagonal elements of the variance and covariance matrix.

3.4.2 MMSIA(k,0)

The model, according to different notations, is also known as MMSI(k), MMSI(k)-VAR(0) or MSIVAR(k,0).

(16) $\mathbf{r_t} = \boldsymbol{\mu}_{\mathbf{S_t}} + \boldsymbol{\varepsilon}_\mathbf{t}$ where $\boldsymbol{\varepsilon}_\mathbf{t} = [\varepsilon_{1\mathrm{t}} \ldots \varepsilon_{\mathrm{nt}}]' \sim N(\mathbf{0}, \boldsymbol{\Sigma})$

The only regime-dependent term is the intercept which is driven by a Markov state variable while the autoregressive coefficients matrix and the covariance matrix are static. Only two types of contagion effects are present in this model configuration: a static and simultaneous one through the off-diagonal elements of the variance and covariance matrix and a dynamic and nonlinear one through the fact that the regime variables that drives the process of all variables are common to all variables.

Also for this model exists a diagonal covariance matrix version which is practically

equivalent to a model that consists of N (number of asset) independent univariate homoskedastic Markov regime switching Normal distributions in which the regime switching dynamic is unique and common for all the asset returns functions. In this case the total log-likelihood of the diagonal MMSIA(k,0) model is equal to the sum of the individually (one for each asset) maximized log-likelihood function.

3.4.3 MMSIAH(k,0)

The model, according to different notations, is also known as MMSIH(k), MMSIH(k)-VAR(0) or MSIVARH(k,0).

(17) $\mathbf{r_t} = \boldsymbol{\mu}_{\mathbf{S_t}} + \boldsymbol{\varepsilon}_\mathbf{t}$ where $\boldsymbol{\varepsilon}_\mathbf{t} = [\varepsilon_{1t} \dots \varepsilon_{nt}]' \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{S_t}})$

The model differs from the last model configuration described because here a regime-dependent variance and covariance matrix which add a dynamic and simultaneous contagious effect through the off-diagonal elements of the variance and covariance matrix is present. Again, also for this model exists a diagonal covariance matrix version which is practically equivalent to a model that consists of N (number of asset) independent univariate Markov regime switching Normal distributions in which the regime switching dynamic is unique and common for all the asset returns functions. In this case the total log-likelihood of the diagonal MMSIA(k,0) model is equal to the sum of the individually (one for each asset) maximized log-likelihood function.

3.4.4 MMSIA(k,p)

The model, according to different notations, is also known as MMSIA(k,p), MMSIA(k)-VAR(p) or MSIVAR(k,p).

(18) $\mathbf{r_t} = \boldsymbol{\mu}_{\mathbf{S_t}} + \sum_{j=1}^{\mathrm{p}} \boldsymbol{A}_{\boldsymbol{j},\mathbf{S_t}} \mathbf{r}_{\mathbf{t}-\boldsymbol{j}} + \boldsymbol{\varepsilon}_\mathbf{t}$

where $\boldsymbol{\varepsilon}_\mathbf{t} = [\varepsilon_{1t} \dots \varepsilon_{nt}]' \sim N(\mathbf{0}, \boldsymbol{\Sigma})$

This is a homoskedastic model configuration in which a dynamic and linear contagious effect through the VAR coefficients, and a dynamic and nonlinear one through the shared regime driven variables, are generated by the regime-

dependent intercept and autoregressive coefficients matrix terms. Again, equivalently to the first three models illustrated before, for this model exists a diagonal covariance matrix version. In this version the covariance between the residuals of the different equations are not allowed thus there is no evidence of a simultaneous contagious effect through the off-diagonal elements of the variance and covariance matrix.

3.4.5 MMSIH(k,p)

The model, according to different notations, is also known as MMSIH(k)-VAR(p).

(19) $$\mathbf{r_t} = \boldsymbol{\mu}_{\mathbf{S_t}} + \sum_{j=1}^{p} \boldsymbol{A_j}\,\mathbf{r_{t-\mathit{j}}} + \boldsymbol{\varepsilon}_\mathbf{t}$$

where $\boldsymbol{\varepsilon}_\mathbf{t} = [\varepsilon_{1t} \dots \varepsilon_{nt}]' \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{S_t}})$. The model is a special case of the MSVAR(k,p) in which while intercepts and the covariance matrix are regime-dependent, the VAR(p) autoregressive coefficients are not. This model configuration shows a dynamic and simultaneous contagious effect through the off-diagonal elements of the variance and covariance matrix, a static and linear one through the VAR coefficients and a dynamic and nonlinear one through the fact that the regime variables that drives the process of all variables is common to all variables. Likewise the last model illustrated for this model exists a diagonal covariance matrix version characterized by the fact that the covariance between the residuals of the different equations are not allowed thus there is no evidence of a simultaneous contagious effect through the off-diagonal elements of the variance and covariance matrix.

3.4.6 MMSI(k,p)

The model, according to different notations, is also known as MMSI(k)-VAR(p).

(20) $$\mathbf{r_t} = \boldsymbol{\mu}_{\mathbf{S_t}} + \sum_{j=1}^{p} \boldsymbol{A_j}\,\mathbf{r_{t-\mathit{j}}} + \boldsymbol{\varepsilon}_\mathbf{t}$$

where $\boldsymbol{\varepsilon}_\mathbf{t} = [\varepsilon_{1t} \dots \varepsilon_{nt}]' \sim N(\mathbf{0}, \boldsymbol{\Sigma})$. The model is a special case of the MSVAR(k,p) in which while intercepts are regime-dependent, the VAR(p) autoregressive coefficients and the covariance matrix are not, then the model is homoskedastic.

This model configuration shows a static and simultaneous contagious effect through the off-diagonal elements of the variance and covariance matrix, a static and linear one through the VAR coefficients and a dynamic and nonlinear one through the fact that the regime variables that drives the process of all variables is common to all variables. In the same way as for all the model previously illustrated there is a less sophisticated version of the model characterized by a diagonal covariance matrix, thus the covariance between the residuals of the different equations are not allowed and for this reason in this version of the model there is no evidence of a simultaneous contagious effect through the off-diagonal elements of the variance and covariance matrix.

3.4.7 Multivariate restricted (common underlying Markov chain) MSVAR(K,1) model

Guidolin[9] proposed a method to obtain a multivariate MSVAR(k,1) model starting from the estimation of single univariate models, one for each asset class, and making assumption on their underlying Markov chains. Each univariate model is a Markov Switching first-order autoregressive model MSARH(k,1). The intercept, the autoregressive coefficient, the regression coefficients and the variance of the shock are regime-dependent. The model takes the following form:

(21) $$\mathrm{r}_{i,\mathrm{t}} = \mu_{i\mathrm{S_t}} + \phi_{\mathrm{S_t}}^{i,i}\mathrm{r}_{i,\mathrm{t}-1} + \sum_{j=1}^{J}\phi_{\mathrm{S_t}}^{i,j}\,\mathrm{r}_{j,\mathrm{t}-1} + \varepsilon_{i,\mathrm{t}} \quad \forall\, i \neq j$$

where $\varepsilon_{i,\mathrm{t}} \sim N(0, \sigma_{i,\mathrm{S_t}}^2)$, $\mathrm{r}_{j,\mathrm{t}-1}$ is the return at time *t-1* on a generic asset or predictor *j*, $\phi_{\mathrm{S_t}}^{i,i}$ is the autoregressive coefficient in state $\mathrm{S_t}$ and $\phi_{\mathrm{S_t}}^{i,j}$ is the regression coefficient of the return of asset *i* on the return of asset *j* in state $\mathrm{S_t}$.

Once *n* univariate MSARH(k,1) models have been estimated, when *n* represent the number of assets, is possible to estimate a multivariate model which is a restricted version of the *n* univariate models. The restriction imposes $\mathrm{S_t} = \mathrm{S}_{i,\mathrm{t}}$ for $i = 1, \dots, n$, in words a unique Markov chain is assumed to drive simultaneously the

[9] Massimo Guidolin "Modelling, Estimating and Forecasting Financial Data under Regime (Markov) Switching" download at "http://didattica.unibocconi.it/mypage/dwload.php?nomefile=Lecture_7_-_Markov_Switching_Models20130520235704.pdf"

regime switching dynamics of all the asset classes. This model may be obtained from a set of *n* univariate models, one for each of the *n* asset returns, when the means, the variance parameters, the autoregressive coefficients and the regression coefficients in the multivariate model are set to be identical to those of the univariate models. The restriction $\mathrm{S_t} = \mathrm{S}_{i,\mathrm{t}}$ for $i = 1, \dots, n$ implies that exists a unique transition matrix and therefore the transition probabilities from one regime to another are identical for all the asset returns processes. For instance if has been chosen a two states model the restriction implies that $p_{11} = p_{i,11}$ and $p_{22} = p_{i,22}$ for $i = 1, \dots, n$. It is important to notice that the model is a multivariate restricted version of *n* univariate models only when the simultaneous covariance coefficients are restricted to be zero, this impose a diagonal structure of the covariance matrix, for this reason the total log-likelihood for the multivariate restricted model is equal to the sum of the log-likelihood of the *n* univariate models individually maximized. A likelihood ratio test is proposed by the author to assess the null hypothesis that the restriction cannot be rejected based on the available data. The model proposed by Guidolin in this framework, for the reason explained just before, is characterized by the fact that the covariance coefficients in all regimes are restricted to be zero, i.e., the only source of correlation in the system is the fact that the same Markov state variables drive simultaneously the switches in all the time series. Given that in the multivariate model the parameters of each asset class are constrained to assume the value estimated from the corresponding single univariate model, in which context covariance terms cannot be estimated, thus the covariance matrix of the multivariate model is also constrained to be diagonal.

The model can be represented in the following way (assuming there are only three assets $\mathrm{r_{i,t}}$ and a predictor $\mathrm{z_t} = \mathrm{r_{4,t}}$, thus *n*=4):

(22)

$$\mathrm{r_{1,t}} = \mu_{1\mathrm{S_t}} + \phi_{\mathrm{S_t}}^{1,1}\mathrm{r_{1,t-1}} + \sum_{j=1}^{4} \phi_{\mathrm{S_t}}^{1,j}\,\mathrm{r}_{j,\mathrm{t-1}} + \varepsilon_{1,\mathrm{t}} \quad \forall\, j \neq 1$$

$$\mathrm{r}_{2,\mathrm{t}} = \mu_{2\mathrm{S_t}} + \phi_{\mathrm{S_t}}^{2,2}\mathrm{r}_{2,\mathrm{t}-1} + \sum_{j=1}^{4} \phi_{\mathrm{S_t}}^{2,j}\, \mathrm{r}_{j,\mathrm{t}-1} + \varepsilon_{2,\mathrm{t}} \quad \forall\, j \neq 2$$

$$\mathrm{r}_{3,\mathrm{t}} = \mu_{3\mathrm{S_t}} + \phi_{\mathrm{S_t}}^{3,3}\mathrm{r}_{3,\mathrm{t}-1} + \sum_{j=1}^{4} \phi_{\mathrm{S_t}}^{3,j}\, \mathrm{r}_{j,\mathrm{t}-1} + \varepsilon_{3,\mathrm{t}} \quad \forall\, j \neq 3$$

$$\mathrm{z_t} = \mathrm{r}_{4,\mathrm{t}} = \mu_{4\mathrm{S_t}} + \phi_{\mathrm{S_t}}^{4,4}\mathrm{r}_{4,\mathrm{t}-1} + \sum_{j=1}^{4} \phi_{\mathrm{S_t}}^{4,j}\, \mathrm{r}_{j,\mathrm{t}-1} + \varepsilon_{4,\mathrm{t}} \quad \forall\, j \neq 4$$

where

$$\begin{bmatrix} \varepsilon_{1,\mathrm{t}} \\ \varepsilon_{2,\mathrm{t}} \\ \varepsilon_{3,\mathrm{t}} \\ \varepsilon_{4,\mathrm{t}} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{1,1,\mathrm{S_t}}^2 & 0 & 0 & 0 \\ 0 & \sigma_{2,2,\mathrm{S_t}}^2 & 0 & 0 \\ 0 & 0 & \sigma_{3,3,\mathrm{S_t}}^2 & 0 \\ 0 & 0 & 0 & \sigma_{4,4,\mathrm{S_t}}^2 \end{bmatrix} \right)$$

is the multivariate normal distribution of the shocks at time *t*, $\mu_{i\mathrm{S_t}}$ is the intercept term of the asset *i*, $\mathrm{r}_{j,\mathrm{t}-1}$ is the return at time *t-1* on a generic asset ($\mathrm{z_t} = \mathrm{r}_{4,\mathrm{t}}$ is the predictor), $\phi_{\mathrm{S_t}}^{i,i}$ is the autoregressive coefficient in state $\mathrm{S_t}$ and $\phi_{\mathrm{S_t}}^{i,j}$ is the regression coefficient of the return of asset *i* on the return of asset *j* in state $\mathrm{S_t}$.

The same model can be represented in matricial form in the following way:

(23) $$\boldsymbol{y}_{\mathbf{t}} = \boldsymbol{\mu}_{\mathbf{S_t}} + \boldsymbol{A}_{\boldsymbol{j},\mathbf{S_t}}\boldsymbol{y}_{\mathbf{t}-\boldsymbol{j}} + \boldsymbol{\varepsilon}_{\mathbf{t}}$$

where

$$\begin{bmatrix} \varepsilon_{1,\mathrm{t}} \\ \varepsilon_{2,\mathrm{t}} \\ \varepsilon_{3,\mathrm{t}} \\ \varepsilon_{4,\mathrm{t}} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{1,1,\mathrm{S_t}}^2 & 0 & 0 & 0 \\ 0 & \sigma_{2,2,\mathrm{S_t}}^2 & 0 & 0 \\ 0 & 0 & \sigma_{3,3,\mathrm{S_t}}^2 & 0 \\ 0 & 0 & 0 & \sigma_{4,4,\mathrm{S_t}}^2 \end{bmatrix} \right)$$

is the multivariate normal distribution of the shocks at time *t*, $\boldsymbol{y}_{\mathbf{t}} = (\mathbf{r}_{\mathbf{t}}' \;\; \mathbf{z}_{\mathbf{t}}')'$ is a *(3 + 1) x* 1, $\boldsymbol{\mu}_{\mathbf{S_t}} = [\mu_{1\mathrm{S_t}} \ldots \mu_{4\mathrm{S_t}}]'$ is the intercept vector for $\boldsymbol{y}_{\mathbf{t}}$ in state $\mathrm{S_t}$ and $\boldsymbol{A}_{\boldsymbol{j},\mathbf{S_t}}$is a *(3 + 1) x (3 + 1)* matrix of autoregressive and regression coefficients.

The following table provide a recap of the different notation commonly used for each model configuration previously described in this paragraph.

Table 5 Commonly used alternative notation for MSVAR family models

	alternative notations			
model specification	MMSI(k)	MMSIA(k,0)	MMSI(k)-VAR(0)	MSIVAR(k,0)
		MMSIA(k,p)	MMSIA(k)-VAR(p)	MSIVAR(k,p)
		MMSIH(k,p)	MMSIH(k)-VAR(p)	
		MMSI(k,p)	MMSI(k)-VAR(p)	
	MMSIH(k)	MMSIAH(k,0)	MMSIH(k)-VAR(0)	MSIVARH(k,0)
		MMSIAH(k,p)	MMSIAH(k)-VAR(p)	MSIVARH(k,p)

3.5 Choice of model specification

The objective of the model estimation procedure is to find the best fitting model and the number of regimes is part of the fit.

Suppose you have a general model parameterized by θ and that you want to test a set of *r* restriction that transforms θ into a sub-vector θ^{r} of dimension *r*. Once both model have been estimated by MLE the maximized log-likelihoods $\ell(\theta)$ and $\ell(\theta^{R})$ are available and the log-likelihood test statistic is $\mathrm{LRT} = 2[\ell(\theta) - \ell(\theta^{R})] \sim \chi_r^2$. Unfortunately is not possible to use the likelihood ratio test , in fact the usual regularity conditions fail because under the null hypothesis, some of the parameters of the model would be unidentified (they are called “nuisances”). For example, if there is really only one regime, the ML estimate of the transition probability matrix does not converge to a well-defined value, meaning that the likelihood ratio test does not have the usual Chi-squared limiting distribution. To interpret a LRT statistic one instead needs to appeal to a simulation methods.

A practical alternative consists of using information criteria that are essentially penalized log-likelihood tests in which the trade-off between fit and parsimony is quantified. The goal is to minimize the value of the information criteria. In general there are two ways to reduce the value of the information criteria: increase the maximized log-likelihood or reduce the number of parameters.

A range of values for the number of regimes is considered $(\mathrm{k} = 1,2,3, \dots)$, doing so very parsimonious as well as heavily parameterized models are covered. To select among the regimes specifications, Akaike (AIC), Schwartz (SIC) or Bayesian (BIC) and Hannan-Quinn (HQC) are usually considered. Unlike formal hypothesis tests which are subject to nuisance parameter problems, these criteria do not, however, provide rigorous tests for the presence of regimes. The AIC tends to suggest the selection of overparameterized models according to Fenton and Gallant (1996), the BIC on the contrary tends to favors small models while Hannan-Quinn is usually in an intermediate position when compared to BIC and AIC. The three information criteria formula are here illustrated:

(24) $AIC = 2 \cdot k - 2 \cdot \ln(logLikelihood_{maximized})$

(25) $BIC = k \cdot \ln(n) - 2 \cdot \ln(logLikelihood_{maximized})$

(26) $HQC = 2 \cdot k \cdot \ln(\ln(n)) - 2 \cdot \ln(logLikelihood_{maximized})$

where *k* stands for the number of free parameters to be estimated, *n* stands for the number of observations while $logLikelihood_{maximized}$is the maximized value of the log-Likelihood for the estimated model.

3.5.1 Alternative models estimation results

In this subparagraph I am going to describe the method I adopted to estimate the model and the results of the estimation procedure. Firstly I started by estimating a range of MSVAR model commonly used by many authors in the relevance literature, and a less common restricted version, secondly I applied an information criteria to select the best model according to it. To determine the best model I undertake an extensive specification search estimating models that span a number of regimes from 1 to 4 and autoregressive order from 1 to 2. This approach covers very parsimonious as well as heavily parameterized models. For the estimation I employed the Marcelo Perlin's MS_Regress, a MATLAB toolbox[10] (October 30, 2014 versions) specially designed for the estimation and simulation of markow regime switching model. The package allows the user to estimate a

[10] Perlin, M. (2014) MS Regress - The MATLAB Package for Markov Regime Switching Models. Available at SSRN: http://ssrn.com/abstract=1714016 or http://dx.doi.org/10.2139/ssrn.1714016

large number of different Markov switching specifications, without any change in the original code. I wrote my own script to manage the estimation routine of several models in a row, including the composition of the input structure for each model into the package's notation , the filling of the package fitting function and the options to feed it, the parsing of the output and the storing of the estimation results and charts. In this Matlab package all of the models are estimated using maximum likelihood. Results of the model specification analysis are presented in Table 6.

Table 6 Estimation results

Model	**Number of parameters (saturation ratio)**	**Log-likelihood**	**AIC**	**BIC**	**Hannan-Quinn**
MMSIAH(k,p)					
MMSIAH(2,1)	62 (34.84)	-12383.33	-12383.33	-12031.30	-12279.27
MMSIAH(3,1)	**96 (22.5)**	**-12996.57**	**-12996.57**	**-12451.49**	**-12835.44**
MMSIAH(3,2)	144 (15)	-12740.35	-12740.35	-11922.74	-12498.66
MMSIAH(4,1)	132 (16.36)	-12575.36	-12575.36	-11825.88	-12353.81
MMSIAH(4,2)	196 (11.02)	-12528.85	-12528.85	-11415.99	-12199.88
MMSIAH(k,0)					
MMSIAH(2,0)	30 (72)	4796.60	-9533.21	-9362.87	-9482.85
MMSIAH(3,0)	48 (45)	5374.93	-10653.85	-10381.32	-10573.29
MMSIAH(4,0)	68 (31.76)	5373.71	-10611.42	-10225.33	-10497.29
MMSIAH(k,p) diagonal covariance matrix					
MMSIAH(2,1) diagonal covariance matrix	50 (43.2)	-11786.48	-11786.48	-11502.59	-11702.56
MMSIAH(3,1) diagonal covariance matrix	78 (27.69)	-12139.21	-12139.21	-11696.33	-12008.29
MMSIAH(3,2) diagonal covariance matrix	126 (17.14)	-12093.47	-12093.47	-11378.05	-11881.98
MMSIAH(4,1) diagonal covariance matrix	168 (12.86)	-12047.66	-12047.66	-11434.45	-11866.39
MMSIAH(4,2) diagonal covariance matrix	172 (12.56)	-11957.77	-11957.77	-10981.18	-11669.08
MMSIAH(k,0) diagonal covariance matrix model					
MMSIAH(2,0) diagonal	18 (120)	4836.79	-9637.57	-9535.37	-9607.36

covariance matrix					
MMSIAH(3,0) diagonal covariance matrix	30 (72)	4914.18	-9768.36	-9598.02	-9718.00
MMSIAH(4,0) diagonal covariance matrix	44 (49.09)	5119.76	-10151.52	-9901.69	-10077.67

The previous table contains the number of estimated parameters, the saturation ratio value, the maximized value of the log-Likelihood and three information criteria values for each model configuration estimated. I decided to use the Hannan-Quinn as model selection method since it holds the middle ground between the Akaike and Schwartz information criteria. The Hannan-Quinn information criteria supports the MMSIAH(3,1) model; I therefore would have settled on a three states model with first order autoregressive components.

Estimating a richly parameterized model in which the number of parameters is so large that the saturation ratio (i.e., the number of observations available to estimate each parameter, on average) is below 20, leads to encounter difficulties at obtaining reliable parameters estimates. A common rule of thumb proposes that nonlinear estimation results based on saturation ratios less than to 20 ought to be taken with great caution.

3.5.2 Multivariate restricted MSVAR(K,1) models estimation results

Given that the preferred model according to the Hannan-Quinn information ratio is characterized to have a low saturation ratio equal to 22.5 and that the two next most supported models, a MMSIAH(3,2) and a MMSIAH(4,1), have saturation ratios respectively of 15 and 16.36, and given that a high number of estimated parameters fail to be statistically significant - probably because of the bad quality of the estimation caused by the high number of free parameters - I preferred to perform the estimation of an additional class of MSVAR family models, namely the multivariate restricted MSVAR(k,1) model. This class of model configuration has been described previously in the subparagraph 3.4.7. I proceeded estimating three MSARH(k,p) models for each asset class. From the adoption of the Hannan-Quinn information ratio emerged that, for three out of four asset classes, a 2 states first order autoregressive model is sufficient to capture and described the asset returns

dynamic. The only times series for which there is evidence that a 2 states model is too parsimonious is the div_y, indeed a 3 states model is supported by the data. Once the 4 univariate MSARH(k,1) models are estimated, and the free estimated parameters available, is possible to proceed to estimate the restricted multivariate MSVAR(k,1) models conditionally to the parameters values previously obtain by the estimation of the 4 univariate MSARH(k,1). The restriction imposes $\mathrm{S_t} = \mathrm{S}_{i,\mathrm{t}}$ for $i = 1, \dots, n$, which means that a unique Markov chain is assumed to drive simultaneously all the regime switching dynamics of all the asset classes time series which implies that each asset class in a MSVAR(k,1) multivariate model is assumed to have a regime switching dynamic characterized by a number of states identical to the ones of all the other asset classes. Lastly, I estimated a multivariate model in which the restriction imposes all the asset classes time series to be explained either by a unique 2 states underlying Markov chain, namely a MSVAR(2,1) or by a unique 3 states underlying Markov chain, namely a MSVAR(3,1) or by an additional 4 sates model called MSVAR(4,1). All the estimation results are presented in the following tables, the bold Hannan-Quinn information ratio value indicates the selected univariate model for each asset class and for the multivariate restricted model.

Table 7 Univariate models estimation results

Model (k,p)	**Number of parameters (saturation ratio)**	**Log-likelihood**	**AIC**	**BIC**	**Hannan-Quinn**
variable: lo20					
MSARH(2,1)	14 (38.6)	760.92	-1,493.84	-1,433.79	**-1,470.35**
MSARH(3,1)	24 (22.5)	768.72	-1,489.43	-1,386.48	-1,449.17
MSARH(4,1)	36 (15)	782.98	-1,493.97	-1,339.54	-1,433.57
variable: hi20					
MSARH(2,1)	14 (38.6)	953.29	-1,878.58	-1,818.52	**-1,855.09**
MSARH(3,1)	24 (22.5)	963.31	-1,878.61	-1,775.66	-1,838.35
MSARH(4,1)	36 (15)	968.64	-1,865.29	-1,710.86	-1,804.89
variable: tbond					
MSARH(2,1)	14 (38.6)	1,411.50	-2,795.01	-2,734.95	**-2,771.52**
MSARH(3,1)	24 (22.5)	1,416.00	-2,784.00	-2,681.05	-2,743.73

MSARH(4,1)	36 (15)	1,419.09	-2,766.17	-2,611.74	-2,705.77
variable: div_y					
MSARH(2,1)	14 (38.6)	2,951.94	-5,875.87	-5,815.82	-5,852.38
MSARH(3,1)	24 (22.5)	3,017.29	-5,986.58	-5,883.63	**-5,946.32**
MSARH(4,1)	36 (15)	2,979.60	-5,887.19	-5,732.76	-5,826.79

Table 8 Multivariate restricted models estimation results

Model	Number of parameters (saturation ratio)	Log-likelihood	AIC	BIC	Hannan-Quinn
MSVAR(2,1) restricted	4 (135)	6,066.08	-12,116.1	-12,070.7	**-12,102.7**
MSVAR(3,1) restricted	9 (60)	6,050.59	-12,075.1	-12,001.3	-12,053.3
MSVAR(4,1) restricted	16 (33.8)	5343.32	-10654.64	-10563.80	-10627.79

3.6 Selected model estimates

According to the Hannan-Quinn information criteria, the multivariate restricted model most supported by the data is the 2 regimes restricted first order autoregressive MSVAR(2,1). The same conclusion would have been reached even if either AIC or BIC information criteria had been used instead of the Hannan-Quinn information criteria.

Table 9 shows parameters estimates (pvalues are reported in parentheses) for the regime switching restricted MSVAR(2,1) model:

(27) $\boldsymbol{y}_{\mathbf{t}} = \boldsymbol{\mu}_{\mathbf{S_t}} + \boldsymbol{A}_{\boldsymbol{j},\mathbf{S_t}}\boldsymbol{y}_{\mathbf{t-1}} + \boldsymbol{\varepsilon}_{\mathbf{t}}$

where $\boldsymbol{y}_{\mathbf{t}} = (\mathbf{r}_{\mathbf{t}}' \ \mathbf{z}_{\mathbf{t}}')'$ is a *(3 + 1) x* 1 vector of excess returns, more precisely lo20, hi20, tbond and div_y, and $\boldsymbol{\mu}_{\mathbf{S_t}} = [\mu_{1\mathrm{S_t}} \dots \mu_{4\mathrm{S_t}}]'$ is the *4 x 1* vector of intercept terms of $\boldsymbol{y}_{\mathbf{t}}$ in state $\mathrm{S_t}$ while $\boldsymbol{A}_{\boldsymbol{j},\mathbf{S_t}}$ is the *(3 + 1) x (3 + 1)* matrix of autoregressive and regression coefficients associated with lag 1 in state $\mathrm{S_t}$ and $\boldsymbol{\varepsilon}_{\mathbf{t}} = [\varepsilon_{1,\mathrm{t}} \ \varepsilon_{2,\mathrm{t}} \ \varepsilon_{3,\mathrm{t}} \ \varepsilon_{4,\mathrm{t}}]' \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{\Sigma}_{S_t})$.

Table 9 restricted MSVAR(2,1) parameter estimates

multivariate restricted MSVAR(2,1)				
	lo20	hi20	tbond	div_y
1. Intercept term				
Regime 1	-0.0032	0.0081	0.0026	0.0004
Regime 2	0.0085	0.0043	0.0018	-0.0009
2. VAR(1) Matrix				
Regime 1				
lo20	0.3542	-0.0245	-0.0335	-0.0002
hi20	-0.1371	-0.1015	0.0047	-0.0138
tbond	0.3931	0.3656	0.3021	-0.0058
div_y	0.2568	0.0071	-0.0476	0.9875
Regime 2				
lo20	0.0664	0.0116	-0.0813	0.0005
hi20	0.2824	0.1214	-0.0255	0.0109
tbond	-0.1335	0.1470	0.3673	0.0164
div_y	-0.0936	-0.1509	0.0040	1.0234
3. Covariance Matrix				
Regime 1				
lo20	0.001764			
hi20	0.000000	0.000955		
tbond	0.000000	0.000000	0.000215	
div_y	0.000000	0.000000	0.000000	0.000001
Regime 2				
lo20	0.007056			
hi20	0.000000	0.003310		
tbond	0.000000	0.000000	0.001026	
div_y	0.000000	0.000000	0.000000	0.000016
4. Transition Probabilities	*Regime 1*	*Regime 2*		
Regime 1	0.8988**** (0.0000)	0.4873**** (0.0000)		
Regime 2	0.1012**** (0.0000)	0.5127**** (0.0000)		

**** denotes significance at 0.01%

As it can be seen from Table 9 the model has 64 parameters; 8 intercept terms, 32 regression and autoregressive coefficients, 20 covariance and variance terms (even though the covariance terms are set to 0 and have not been estimated) and 4 regimes transition probabilities. The 4 regimes transition probabilities are the only free parameters that have been estimated in the multivariate restricted model MSVAR(2,1) estimation (their pvalues are reported in parentheses), the totality of

the other parameters in the estimation procedure have been set to be equal to the values estimated from the corresponding univariate MSARH(2,1) models. Table 10 reports the parameter estimates for each one of the four univariate models.

Table 10 single univariate MSARH(2,1) parameter estimates

single univariate MSARH(2,1) models				
	lo20	hi20	tbond	div_y
1. Intercept term				
Regime 1	-0.0032 (0.6758)	0.0081*** (0.0003)	0.0026 (0.2427)	0.0004**** (0.0000)
Regime 2	0.0085 (0.5602)	0.0043 (0.7199)	0.0018 (0.5738)	-0.0009 (0.8432)
2. VAR(1) Matrix				
Regime 1				
lo20	0.3542**** (0.0000)	-0.0245 (0.6726)	-0.0335 (0.1747)	-0.0002 (0.6582)
hi20	-0.1371 (0.1917)	-0.1015 (0.3295)	0.0047 (0.9060)	-0.0138**** (0.0000)
tbond	0.3931** (0.0025)	0.3656** (0.0048)	0.3021**** (0.0000)	-0.0058** (0.0048)
div_y	0.2568 (0.2551)	0.0071 (0.9415)	-0.0476 (0.5279)	0.9875**** (0.0000)
Regime 2				
lo20	0.0664 (0.5200)	0.0116 (0.8995)	-0.0813 (0.6747)	0.0005 (0.9464)
hi20	0.2824 (0.0882)	0.1214 (0.3805)	-0.0255 (0.9199)	0.0109* (0.0326)
tbond	-0.1335 (0.6511)	0.147 (0.5035)	0.3673*** (0.0005)	0.0164 (0.5724)
div_y	-0.0936 (0.8356)	-0.1509 (0.6978)	0.004 (0.9223)	1.0234**** (0.0000)
3. Variance				
Regime 1	0.001764**** (0.0000)	0.000955**** (0.0000)	0.000215**** (0.0000)	0.000001** (0.0024)
Regime 2	0.007056**** (0.0000)	0.003310**** (0.0000)	0.001026**** (0.0000)	0.000016 (0.1865)

* denotes significance at 5%, ** at 1%, *** at 0.1%, **** at 0.01%

I voluntary omitted from the previous table the transition probabilities estimates and their relative pvalues because, conversely to the other parameter estimates, they have not been used to set any parameter values in the estimation of the multivariate restricted model. It can be noticed that 20 parameters out of 48

estimated from the single univariate models are statistically significant at the 10% level, the latter conclusion implies that the total number of statistically significant parameters in the multivariate restricted model are 24 considering the 4 statistically significant at the 10% level regimes transition probabilities estimated in this latter multivariate model. Thus it can be concluded that the number of statistically significant parameters at the 10% level of the multivariate restricted model is 24 out of 52; while if we consider the total number of parameters of the model, which includes also not estimated parameters such as the covariance terms set to 0, the number of statistically significant parameters is 28 out of 64. The parameter estimates are summed up in the following model expressions which have been previously explained in the subparagraph 3.4.7.:

(28) $\boldsymbol{y}_{\mathbf{t}} = \boldsymbol{\mu}_{\mathbf{S_t}} + \boldsymbol{A}_{\mathbf{1,S_t}}\boldsymbol{y}_{\mathbf{t-1}} + \boldsymbol{\varepsilon}_{\mathbf{t}}$

Regime 1

$$\begin{bmatrix} lo20_t \\ hi20_t \\ tbond_t \\ div\_y_t \end{bmatrix} = \begin{bmatrix} -0.0032 \\ 0.0081 \\ 0.0026 \\ 0.0004 \end{bmatrix} + \begin{bmatrix} 0.3542 & -0.1371 & 0.3931 & 0.2568 \\ -0.0245 & -0.1015 & 0.3656 & 0.0071 \\ -0.0335 & 0.0047 & 0.3021 & -0.0476 \\ -0.0002 & -0.0138 & -0.0058 & 0.9875 \end{bmatrix} \begin{bmatrix} lo20_{t-1} \\ hi20_{t-1} \\ tbond_{t-1} \\ div\_y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{lo20,\mathrm{t}} \\ \varepsilon_{hi20,\mathrm{t}} \\ \varepsilon_{tbond,\mathrm{t}} \\ \varepsilon_{div\_y,\mathrm{t}} \end{bmatrix}$$

$$\begin{bmatrix} \varepsilon_{lo20,t} \\ \varepsilon_{hi20,t} \\ \varepsilon_{tbond,t} \\ \varepsilon_{div\_y,t} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.001764 & 0 & 0 & 0 \\ 0 & 0.000955 & 0 & 0 \\ 0 & 0 & 0.000215 & 0 \\ 0 & 0 & 0 & 0.000001 \end{bmatrix} \right)$$

Regime 2

$$\begin{bmatrix} lo20_t \\ hi20_t \\ tbond_t \\ div\_y_t \end{bmatrix} = \begin{bmatrix} 0.0085 \\ 0.0043 \\ 0.0018 \\ -0.0009 \end{bmatrix} + \begin{bmatrix} 0.0664 & 0.2824 & -0.1335 & -0.0936 \\ 0.0116 & 0.1214 & 0.1470 & -0.1509 \\ -0.0813 & -0.0255 & 0.3673 & 0.0040 \\ 0.0005 & 0.0109 & 0.0164 & 1.0234 \end{bmatrix} \begin{bmatrix} lo20_{t-1} \\ hi20_{t-1} \\ tbond_{t-1} \\ div\_y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{lo20,\mathrm{t}} \\ \varepsilon_{hi20,\mathrm{t}} \\ \varepsilon_{tbond,\mathrm{t}} \\ \varepsilon_{div\_y,\mathrm{t}} \end{bmatrix}$$

$$\begin{bmatrix} \varepsilon_{lo20,t} \\ \varepsilon_{hi20,t} \\ \varepsilon_{tbond,t} \\ \varepsilon_{div\_y,t} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.007056 & 0 & 0 & 0 \\ 0 & 0.003310 & 0 & 0 \\ 0 & 0 & 0.001026 & 0 \\ 0 & 0 & 0 & 0.000016 \end{bmatrix} \right)$$

Transition probabilities

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 0.8988 & 0.4873 \\ 0.1012 & 0.5127 \end{bmatrix}$$

Where the term p_{12}indicates the probability of a regime switch from state 2 at time *t-1* to state 1 at time *t*. The process described by the model can be interpreted using the framework in Samuelson (1991) as a momentum process, i.e., a process that is more likely to continue in the same state rather than transition to the other state.

3.7 Model restriction test

At this point of the study once the estimation has been performed I decided to formally test the restriction implied by the model. The restriction, as described previously, imposes $S_t = S_{i,t}$ for $i = 1, \dots, n$, in words a unique Markov chain is assumed to drive simultaneously all the regime switching dynamics of all the asset classes time series. The restriction $S_t = S_{i,t}$ for $i = 1, \dots, n$ implies that exists a unique transition matrix and therefore the transition probabilities from one regime to another are identical for all the asset returns processes. In the two states model I have estimated the restriction imposes that $p_{11} = p_{i,11}$ and $p_{22} = p_{i,22}$ for $i = 1, \dots, n$. As suggested by Guidolin in his work "Modelling, Estimating and Forecasting Financial Data under Regime (Markov) Switching", I formally tests the restriction performing a Likelihood Ratio test. The test compares the maximized log-likelihood values of the restricted case and the unrestricted case. The restricted case is the multivariate restricted MSVAR(2,1) model while the unrestricted case is the general case with potentially separate regime process. The maximized log-likelihood of the former model is exactly the maximized log-likelihood value of the multivariate restricted MSVAR(2,1) model while the maximized log-likelihood of the latter model is equivalent to the sum of the individually (one for each asset class) maximized log-likelihood function. The log-likelihood ratio statistics is computed in the following way:

(29) $LRT = 2 \cdot (LogLikelihood_{Unrestricted} - LogLikelihood_{Restricted})$

where

$LogLikelihood_{Restricted} = LogLikelihood_{MSVAR(2,1)} = 6066.0830$

$LogLikelihood_{Unestricted} =$
$LogLikelihood_{lo20;MSARH(2,1)} + LogLikelihood_{hi20;MSARH(2,1)} +$
$LogLikelihood_{tbond;MSARH(2,1)} + LogLikelihood_{div\_y;MSARH(2,1)} = 760.9213 +$
$953.2881 + 1411.5042 + 2951.9365 = 6077.6501$

Thus the LRT statistics assumes value 23.1343. The LRT statistics is distributed as a Chi-squared distribution with *g* degrees of freedom where *g* is equal to the number of equality restrictions imposed by the restricted model. In this study 8 parameters are set to be identical:

$p_{11} = p_{lo20,11} = p_{hi20,11} = p_{tbond,11} = p_{div\_y,11}$

$p_{22} = p_{lo20,22} = p_{hi20,22} = p_{tbond,22} = p_{div\_y,22}$

thus the LRT statistics is distributed as a Chi-squared distribution with 8 degrees of freedom.

The null hypothesis of the presence of unique Markov chain that is assumed to drive simultaneously all the regime switching dynamics of all the asset classes time series cannot be rejected based on the available data; the LRT statistics of 23.1343 is smaller than the critical value at alpha 0.1% (one chance over a thousand that the test statistics is wrong) which is equal to 26.1245. The pvalue of the LRT statistics is lower than 0.999 and is equal to 0.9968. To sum up it can be said that there is a statistical evidence, one chance out of a thousand to be wrong, that there is the presence of a unique switching dynamics shared by all the asset classes time series and that the multivariate restricted model MSVARH(2,1) specification cannot be rejected based on statistical evidence.

Figure 36 scatter of the state 1 smoothed probabilities estimated from the single univariate markow switching models

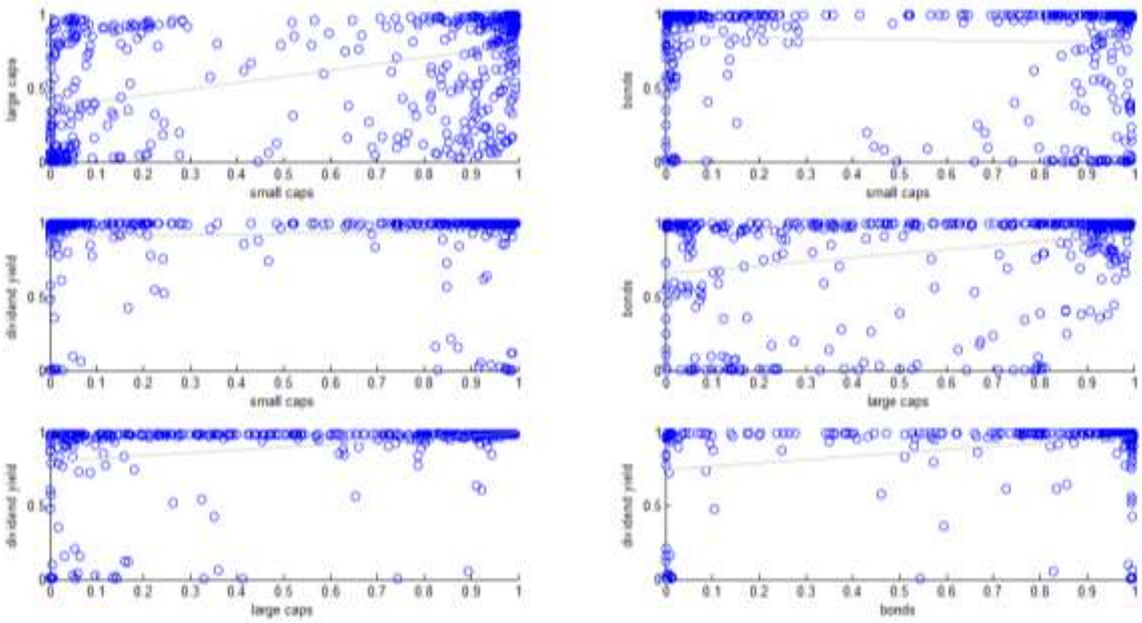


Figure 36 consists of six scatter diagrams each of one represents the relation between the state 1 smoothed probabilities of a pair of returns or dividend yield estimated from single univariate markow switching models. For example the first scatter diagram in the upper-left corner represents the relation between the state 1 smoothed probability of a single univariate markow switching model estimated for the large stocks and the state 1 smoothed probability of a single univariate markow switching model estimated for the small stocks. Overall the predominant relation that emerges from the scatter diagrams seems to be moderately positive except for the pair bonds-small stocks, thus it can be said that, although not so graphically convincing, there is evidence of a positive relation between the state 1 smoothed probabilities of a pair of asset classes or dividend yield estimate from single univariate markow switching models, in other words the greater the smoothed probability of an asset class returns or dividend yield to belong to state 1 the greater the probability of the other asset class returns or dividend yield of the pair to belong to state 1.

At this point I also assessed whether the process is stationary or not. A not regime switching model is considered stationary if the absence of roots outside the unit circle is verified for the matrices of autoregressive coefficients. A generalization of the rule that works for testing stationarity in a not regime switching model, for example a VAR(1), is applicable in the regime switching framework, in fact, Ang

and Bekaert (2002) have showed that formally, it is just sufficient for such a condition to be verified in at least one of the *k* regimes. $\boldsymbol{A}_{\mathbf{1},\mathbf{S_t=1}}$ is the first order matrix of autoregressive and regression coefficients in state 1 while $\boldsymbol{A}_{\mathbf{1},\mathbf{S_t=2}}$ is the state 2 equivalent object. Consider the previously estimated model, MSVAR(2,1), let $\lambda_{i,lo20}, \lambda_{i,hi20}, \lambda_{i,tbond}, \lambda_{i,div\_y}$ be the eigenvalues of the state *i* first order matrix of autoregressive and regression coefficients, $\boldsymbol{A}_{\mathbf{1},\mathbf{S_{t=i}}}$, these eigenvalues solve the characteristic equation $\left|\boldsymbol{A}_{\mathbf{1},\mathbf{S_t=i}} - \lambda \mathbf{I}_{\mathrm{N}}\right| = 0$ then, if the eigenvalues, for just any *i*, are not all equal and smaller than one in modulus thus the MSVAR(2,1) process is stable and given that a stable MSVAR(2,1) implies stationarity then the MSVAR(2,1) process is also stationary. I have calculated and assessed the modulus of the 8 eigenvalues, the results are shown below:

Regime1

$$\begin{bmatrix} \lambda_{1,lo20} \\ \lambda_{1,hi20} \\ \lambda_{1,tbond} \\ \lambda_{1,div\_y} \end{bmatrix} = \begin{bmatrix} 0.3287 + 0.0933i \\ 0.3287 + 0.0933i \\ -0.1035 + 0.0000i \\ 0.9884 + 0.0000i \end{bmatrix} \qquad \left[\begin{vmatrix} \lambda_{1,lo20} \\ \lambda_{1,hi20} \\ \lambda_{1,tbond} \\ \lambda_{1,div\_y} \end{vmatrix}\right] = \begin{bmatrix} 0.3417 \\ 0.3417 \\ 0.1035 \\ 0.9884 \end{bmatrix}$$

Regime2

$$\begin{bmatrix} \lambda_{2,lo20} \\ \lambda_{2,hi20} \\ \lambda_{2,tbond} \\ \lambda_{2,div\_y} \end{bmatrix} = \begin{bmatrix} -0.0312 \\ 0.2752 \\ 0.3125 \\ 1.0221 \end{bmatrix} \qquad \left[\begin{vmatrix} \lambda_{2,lo20} \\ \lambda_{2,hi20} \\ \lambda_{2,tbond} \\ \lambda_{2,div\_y} \end{vmatrix}\right] = \begin{bmatrix} 0.0312 \\ 0.2752 \\ 0.3125 \\ 1.0221 \end{bmatrix}$$

As it can be seen the stability and stationarity of the process in state 1, according to Ang e Bekaert (2002), guarantees the whole process to be stationary.

3.7 Model description

Table 11 unconditional and state conditional asset class and dividend yield means

	lo20	hi20	tbond	div_y
Regime 1	0.004024	0.008032	0.001947	0.023979
Regime 2	0.004980	-0.000730	0.002509	0.035556
unconditional	0.004188	0.006526	0.002044	0.025969

Table 12 unconditional and state conditional covariance matrices adjusted for the regime structure

	lo20	hi20	tbond	div_y
Regime 1				
lo20	0.002300			
hi20	-0.000001	0.001200		
tbond	0.000000	0.000000	0.000297	
div_y	0.000001	-0.000009	0.000001	0.000014
Regime 2				
lo20	0.004478			
hi20	-0.000002	0.002182		
tbond	0.000000	-0.000001	0.000631	
div_y	0.000003	-0.000025	0.000002	0.000042
Unconditional				
lo20	0.002674			
hi20	-0.000001	0.001371		
tbond	0.000000	-0.000001	0.000355	
div_y	0.000002	-0.000014	0.000001	0.000022

Timmerman (2000) derives the moments of a general regime switching process with constant probabilities; the formulas were originally derived for a univariate case, I have adapted them to the multivariate case.

Table 11 shows unconditional and state conditional asset class and dividend yield means calculated as follows:

(30) $\mathrm{E}(\boldsymbol{y}_{\mathbf{t+1}}|\mathrm{S_t}=1;\boldsymbol{\theta})=\overline{\boldsymbol{\mu}_{\mathbf{1}}}=(\mathbf{I_4}-\boldsymbol{A}_{\mathbf{1},\mathbf{1}})^{-1}\boldsymbol{\mu}_{\mathbf{1}}$

(31) $\mathrm{E}(\boldsymbol{y}_{\mathbf{t+1}}|\mathrm{S_t}=2;\boldsymbol{\theta})=\overline{\boldsymbol{\mu}_{\mathbf{2}}}=(\mathbf{I_4}-\boldsymbol{A}_{\mathbf{1},\mathbf{2}})^{-1}\boldsymbol{\mu}_{\mathbf{2}}$

(32) $\mathrm{E}(\boldsymbol{y}_{\mathbf{t+1}}|\boldsymbol{\theta})=\xi_{1\infty}\overline{\boldsymbol{\mu}_{\mathbf{1}}}+(1-\xi_{1\infty})\overline{\boldsymbol{\mu}_{\mathbf{2}}}$

The first two formulas represent the means of a VAR(1) model, in fact conditionally to a current state an MSVAR(1) is equivalent to a VAR(1). $\boldsymbol{A}_{\mathbf{1},\mathbf{1}}$ and $\boldsymbol{A}_{\mathbf{1},\mathbf{2}}$ are *(4 x 4)* matrices of autoregressive and regression coefficients at lag 1, respectively in

state 1 and state 2, $\mathbf{I_4}$ is a *(4 x 4)* identity matrix while $\xi_{1\infty}$ and $\xi_{2\infty}$ are respectively state 1 and state 2 unconditional or ergodic probabilities.

Table 12 shows unconditional and state conditional covariance matrices adjusted for the regime structure calculated as follows:

(33) $Var(\boldsymbol{y_{t+1}}|\mathrm{S_t} = 1; \boldsymbol{\theta}) = \mathrm{p}_{11}\boldsymbol{\Sigma}_1 + \mathrm{p}_{21}\boldsymbol{\Sigma}_2 + \mathrm{p}_{11}\mathrm{p}_{21}(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})^{\boldsymbol{T}}$

(34) $Var(\boldsymbol{y_{t+1}}|\mathrm{S_t} = 2; \boldsymbol{\theta}) = \mathrm{p}_{12}\boldsymbol{\Sigma}_1 + \mathrm{p}_{22}\boldsymbol{\Sigma}_2 + \mathrm{p}_{12}\mathrm{p}_{22}(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})^{\boldsymbol{T}}$

(35) $Var(\boldsymbol{y_{t+1}}|\boldsymbol{\theta}) = \xi_{1\infty}\boldsymbol{\Sigma}_1 + (1 - \xi_{1\infty})\boldsymbol{\Sigma}_2 + \xi_{1\infty}(1 - \xi_{1\infty})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})^{\boldsymbol{T}}$

The first two formulas[11] represents the state conditional covariance matrices, as it can be seen the actual covariance matrices consist of two components. The first component in these formulas are simply weighted averages of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ while the second component takes into account the regime structure, in fact the covariance matrix at time *t+1* depends on the realization of the current regime. The regime structure adds also a jump component to the conditional covariance matrices due to the presence of conditional means that change from one regime to the other. As evidenced by Ang and Timmermann (2012) differences in means across regimes, $\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}}$, enter the higher moments such as variance. In particular, the covariance matrix is not simply the average of the covariance matrices across the two regimes indeed the difference in means also imparts an effect because the switch to a new regime contributes to volatility. Intuitively, the possibility of changing to a new regime with different mean introduces an extra source of risk. Differences in means in addition to differences in covariance matrices can generate persistence in levels and squared values[12], causing volatility persistence observed in many return series. The first formula represents the auto covariance of levels (mean persistence) while the second formula represents auto covariance of squared levels (volatility persistence).

(36) $\mathrm{Cov}(\boldsymbol{y_t}, \boldsymbol{y_{t-1}}) = \xi_{1\infty}(1 - \xi_{2\infty})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})^{\boldsymbol{T}}(\mathrm{p}_{11} + \mathrm{p}_{22} - 1)$

[11] Andrew Ang and Geert Bekaert “How Do Regimes Affect Asset Allocation” (2002) and Massimo Guidolin and Federica Ria “Regime Shifts in Mean-Variance Efficient Frontiers: Some International Evidence” (2010)
[12] Andrew Ang and Allan Timmermann “Regime Changes and Financial Markets” (2012)

(37) $\mathrm{Cov}(\boldsymbol{y}_{\mathbf{t}}, \boldsymbol{y}_{\mathbf{t-1}}) = \xi_{1\infty}(1-\xi_{2\infty})(\overline{\boldsymbol{\mu}_{\mathbf{1}}} - \overline{\boldsymbol{\mu}_{\mathbf{2}}} + \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)(\overline{\boldsymbol{\mu}_{\mathbf{1}}} - \overline{\boldsymbol{\mu}_{\mathbf{2}}} + \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^{\boldsymbol{T}}(\mathrm{p}_{11} + \mathrm{p}_{22} - 1)$

Again differences in means play an important role in generating autocorrelation in first moments, without such differences the autocorrelation will be equal to zero. In contrast, volatility persistence can be induced either by differences in means or by differences in covariance matrices across regimes. In both cases, the persistence tends to be greater the stronger the combined persistence of the regimes, as measured by $(\mathrm{p}_{11} + \mathrm{p}_{22} - 1)$.

The following subparagraph of the paragraph represents an attempt to give an economic interpretation to the model estimated.

3.8.1 Economic interpretation of regimes

Regime 1 is an extremely persistent bull regime characterized by low realized volatility and positive average realized excess returns on all assets. Small stocks and large stocks grow rapidly on average, the average realized excess return on bonds, likewise the average realized dividend yield level, are more modest and relatively lower than in regime 2. The average duration is equal to 9.8855 months and as a result this regime characterizes approximately 85% of the data in the long run. This statistic is the simple empirical probability of being in state 1 computed by dividing the number of draws associated with state 1 by the total number of sample draws (540); a draw is considered to be associated to a state *i* when the corresponding smoothed probability of state 1 is greater than or equal to 0.5%. In this state the intercept term of lo20 and tbond are equal respectively to -0.0032 and 0.0026 but not significantly different from zero while the intercept term of hi20 and the dividend yield are instead statistically different from 0 and are equal respectively to 0.0081 and 0.0004. The volatility of all the asset classes returns and the dividend yield are small and lower than in state 2; small stock returns are the most volatile asset followed by large stocks and bonds. Figure 39 shows that regime 1 capture episodes of the bull market since the mid-sixties, long periods with growing stock prices during the mid-1990s and the late 1990s and the

recent bull market in the mid-2000s. Regime 2 is a not-highly persistent bear regime characterized by high volatility and large and negative average realized excess returns on small and large stocks while the average realized excess return on bonds are significantly positive and relatively larger than in regime 1. The average duration is equal to 2.0523 months and as a result this regime is characterized by approximately 15% of the data in the long run. In this state the intercept terms are all not significantly different from zero. The volatility of all the asset classes returns and the dividend yield are large and higher than in state 1; similarly to state 1 small stock returns are the most volatile asset followed by large stocks and bonds. Figure 39 shows that regime 2 includes two oil shocks in the 1970s, the recession of the early 1980s, the October 1987 stock market crash, the Kuwait invasion in the early 1990s and the ‘Asian flu’ (1996-1998), the dot-com market crash of 2000-2001 (the recent bear market of 2002-2002) and the 2008-2009 financial crisis.

The cross-regime differences between the estimated intercept terms are all smaller than 1% except for lo20 which is equal to 1.1740%, the second larger difference, 0.3768%, belongs to hi20 followed by tbond and div_y ones, equal to 0.0767% and 0.1291% respectively. The cross-regime differences between the estimated volatilities again tend to be larger for lo20 and hi20, respectively equal to 0.5290% and 0.2355% while tbond and div_y have smaller ones equal to 0.0811% and 0.0016%.

Table 13 state conditional time series average levels and volatilities

STATE	STATISTICS	lo20	hi20	tbond	div_y	div_y %change
1	average values	0.010920	0.006673	0.000543	0.030128	0.003674
	volatilites	0.002330	0.001102	0.000268	0.000121	0.000808
2	average values	-0.016294	-0.013784	0.007353	0.036819	-0.021052
	volatilites	0.014520	0.006211	0.001168	0.000241	0.004827

According to Table 13 the larger average realized returns go along with the lower realized volatilities in regime 1 which is considered the bull market state; from the

same table it can be clearly stated that regime 2 is a recession or bear state with high realized volatility and mostly negative average realized returns.

Table 13 shows the state conditional average values and volatilities for small stocks, large stocks, bonds, dividend yield and changes in dividend yield. The values are calculated ex post classifying each period as either state 1 or state 2 based on the estimated smoothed state probabilities; a state 1 (state 2) smoothed probability value greater than 0.5 implicates the classification of that data point as belonging to state 1 (state 2). It follows that the statistics relative to state 2 are calculated from a time series with length 89 whereas the data relative to state 1 are calculated from a time series with length 459. At first glance it can be seen that both the stock returns show a positive average value in state 1 and a negative value in state 1 whereas the average bonds returns are positive in both state similarly to the dividend yield ones, however both the bonds returns and the dividend yield are characterized by higher average values in state 2 than in state 1.

The lo20 and hi20 average levels in state 1 are relative high and above global average level (1.0920% and 0.6673%) while the volatilities are below global average (0.233% and 0.1102%); on the contrary in state 2 the lo20 and hi20 average levels are below global average level (-1.6294% and -1.3784%) and their volatilities are above global average (1.452% and 0.6211%). The tbond time series has a behavior antithetical to that of lo20 and hi20 regarding the average levels in state 1 and 2, in fact in state 1 tbond average level is relatively low and below global average (0.0543%) while in state 2 relative high and above global average (0.7353%); on the contrary the tbond volatility behaves similarly to those of lo20 and hi20 showing lower volatility below global value in state 1 (0.0268%) and higher volatility above global value in state 2 (0.1168%). The dividend yield value is relatively low and below the global mean in state 1 (3.0128%) and the volatility (0.0121%) is also below average, on the contrary in state 2 the posterior average dividend yield value is higher than average (3.6819%) as well as the volatility (0.241%).

I also computed the dividend monthly yield percentage variations and made some considerations on their state conditional average levels and volatilities; I find out

that in state 1 the average dividend yield monthly percentage variation is slightly positive and above the global average (0.3674%) while in state 2 it assumes a dramatic negative value (-2.1052%), similarly the volatility of the dividend yield monthly percentage variation assumes lower value in state 1 than in state 2 (respectively 0.0808% and 0.1481%). The changes in dividend yield present a positive average value in state 1 and a negative average value in state 2, this feature can be explained by the transient nature of the dividend yield (from frequent and long period of low values in state 1 to rare and short high values in state 2) that impose to its changes to be moderately positive in state 1 and significantly negative in state 2. As it can be clearly seen from Table 13, all the asset classes returns, the dividend yield and the changes in dividend yield show larger volatility values in state 2 than in state 1.

Table 14 regimes conditional risk premia

STATE	lo20-hi20	lo20-tbond	hi20-tbond
1	0.004247	0.010377	0.006130
2	-0.002510	-0.023647	-0.021137

Table 14 shows the regimes conditional risk premia calculated as the state conditional differences between the average realized returns on two different asset classes.

As it can be easily seen all the asset classes present big difference in the risk premia among regimes, in fact all the risk premia are positive in state 1 and negative in state 2. State 1 and state 2 identify a size effect in stock returns. In state 1 the average realized return of small stocks exceeds that of large stocks of about 0.4247% per month , while this get reversed in state 2 where the average realized return of large stocks exceeds that of small stocks of about 0.251% per month. It can be seen that the average realized of small stocks and large stocks exceed that of bonds of about 1.0377% and 0.613% respectively in state 1, while in state 2 the situation get reversed with bond returns now exceeding small stocks and large stocks average realized returns of about -2.364% and -2.1137% respectively. The entire risk premia structure is reversed in state 2, this means that

for each pair of asset classes and for each regime, there is a different asset class that beats the other in term of higher average realized return, i.e. in state 1 small stocks beat large stocks and bonds and similarly large stocks beat bonds whereas in state 2 large stock beats small stocks and similarly bonds beat small stocks and large stocks. The considerations above brings to the conclusion that in regime 1 it seems reasonable to invest in equity, whose average realized return outperforms that of bonds, while in regime 2 holding bonds seems to be more profitable. The presence of a size effect brings to the further conclusion that, regarding the stock market, a strategy that invests in small stocks in regime 1 and large stocks in regime 2 seems to be appropriate and profitable.

In this section of the subparagraph some considerations about the liner correlation between returns are provided.

Table 15 regimes conditional linear correlation

STATE	lo20-hi20	lo20-tbond	hi20-tbond
1	0.607525	0.027858	0.177233
2	0.811323	0.268023	0.296864
unconditional	0.718089	0.121093	0.207156

As shown in Table 15 correlations between returns appear to vary substantially across regimes. The realized correlation between large (hi20) and small (lo20) stocks' returns varies from a high of 0.811323 in state 2 to a low of 0.607525 in state1. The correlation between returns on small stocks and bonds varies from a high of 0.268023 in state 2 to a low of 0.027858 in state 1 while the correlation between returns on large stocks and bonds goes from 0.177233 in state 1 to 0.296864 in state 2. The large cross-regime differences in the realized linear correlation coefficients occurs between small stocks and bonds, equal to 0.240164, while the smaller one occurs between large stocks and bonds, equal to 0.119631. Small stock returns have a larger cross-regime difference in realized correlation with bonds returns (0.240164) than large stock returns (0.203798) while large stock returns have it larger with small stock returns (0.203798) than with bonds returns (0.119631). As it can be seen from Table 15 the unconditional

correlations assume values between the two states conditional correlation values, this is consistent with the evidence of regime time-varying correlations found in monthly equity returns by Ang and Chen (2002). The ability of this model to identify a correlation close to 0 between small stocks and bonds returns in state 1 - which the linear model is unable to do – is a sign of the potential value of adopting a regime switching model in the portfolio construction context. As evidenced by recent works by Andersen, Bollerslev, Diebold, and Vega (2004) stock and bond returns move together more than how measured by a single state linear correlation; the reason is that the correlation switches sign across different regimes and may appear spuriously small when averaged across states. The results on realized linear correlation coefficients between returns, conditional to my model states estimates, are consistent with the existing regime switching literature, in fact many studies (e.g., Hong, Tu and Zhou 2007) on asymmetric co-movements between asset returns and market indices suggest that stocks are more likely to move with the market when the market goes down than when it goes up, in fact conditional to the states estimates of my model, all pair-wise realized linear correlation coefficients between returns are structurally higher in state 2 (bear state) than in state 1 (bull state), for instance the average correlation in state 1 is 0.270872 vs. an average of 0.458737 in state 2. As suggested by Jun Tu (2010) Regime-dependent co-movements could be important for examining the economic value of regime switching for portfolio decisions since, thought standard investment theory advises portfolio diversification under pricing model uncertainty, the value of this advice might be questionable if all stocks tend to fall a lot as the market falls in a bear regime.

3.8.2 Regimes and predictability from the dividend yield

Table 16 regime conditional liner correlations between stocks and bonds returns, and dividend yield

STATE	lo20-div_y	hi20-div_y	tbond-div_y
1	0.030009	-0.033811	-0.013665
2	0.002636	0.110058	-0.083438
unconditional	-0.011503	-0.015227	-0.011708

As suggested by Vo and Maurer (2013) the estimated regime switching model can also model the asymmetry in the dividend yield's ability to hedge shocks in the investment opportunity set. As shown in Table 16 each state is characterized by a marked realized correlation structure. During expansion periods both large stock returns and bonds returns tend to move in the opposite direction of the dividend yield, whereas small stock returns offer an hedge against adverse movements of the investment opportunity set since it is characterized to be, although very weakly, negatively correlated with dividend yield. During recession periods both small stock returns and large stock returns co-move with the dividend yield with the latter one showing a considerably higher linear realized correlation coefficient than the former one, conversely the bonds returns offer an hedge against stock returns since it is characterized by a negative, although very weak, realized linear correlation coefficient. Vo and Maurer (2013) also observed that an high correlation between stock returns and the dividend yield implicates a weaker hedge from stocks holdings against adverse movements of the investment opportunity set due to the fact that an unexpected decline in the dividend yield is accompanied by unexpected higher realized stock returns. At the same time however, mean reversion in the dividend yield process lower the investor's expectation about future stock returns. Thus as investment opportunity turn sour, holding a greater amount of stocks represents a more valuable hedge the lower the correlation between stocks and the dividend yield.

Table 17 regime conditional correlations between stocks and bonds returns, and changes in dividend yield

STATE	lo20-Δ%div_y	hi20-Δ%div_y	tbond-Δ%div_y
1	0.591008	0.507782	0.156789
2	0.693650	0.764479	0.241957
unconditional	0.654009	0.649582	0.161971

I also studied the state conditional realized linear correlation coefficients between stock and bond returns, and changes in dividend yield. The results are shown in Table 17. As it can be seen all the asset classes returns show in both regimes

positive realized linear correlation coefficients between them and the changes in dividend yield. A possible explanation for this evidence may be provided by the fact that, as shown in Table 13, in state 1 both the changes in dividend yield and stock returns have positive average values whereas in state 2 they are characterized by negative average values, this reasonably lead to positive realized linear correlation coefficients. The correlation coefficient between changes in dividend yield and bonds returns in state 2 is more challenging to explain and interpret, at first glance, based on Table 13 a negative realized correlation coefficient might be guessed, however the correlation coefficient between changes in dividend yield and bonds returns is weak and its magnitude is not comparable with the correlation coefficients between changes in dividend yield and stock returns. All the considerations regarding the realized linear correlation coefficients provided so far are also supported by the following scatter diagrams.

Figure 37 scatter of the dividend yield changes vs. small, large and bonds returns

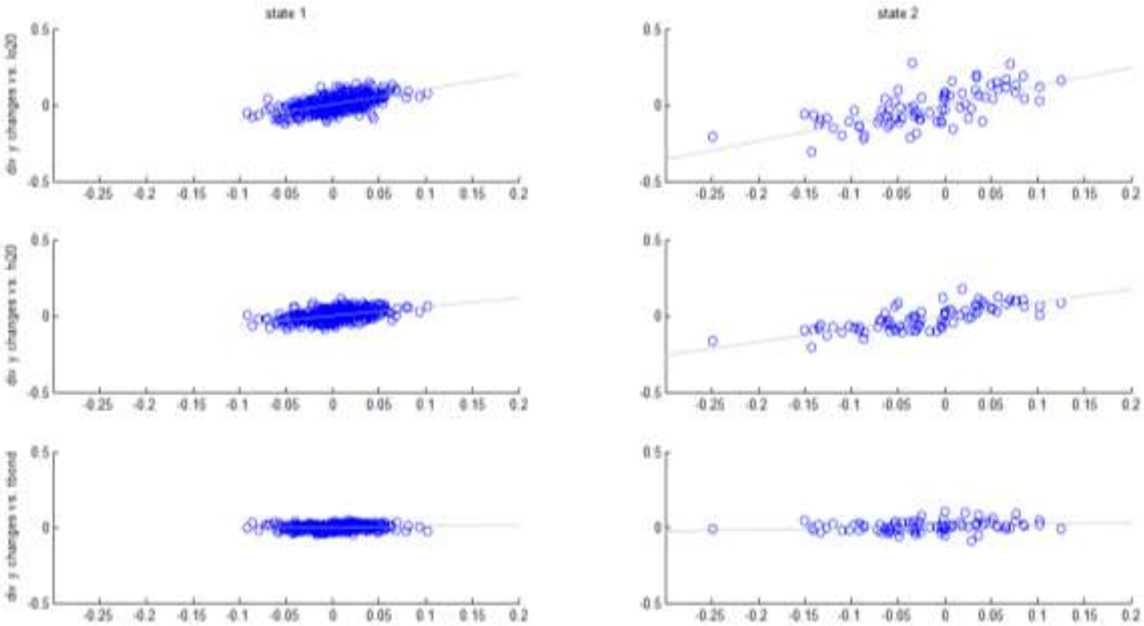
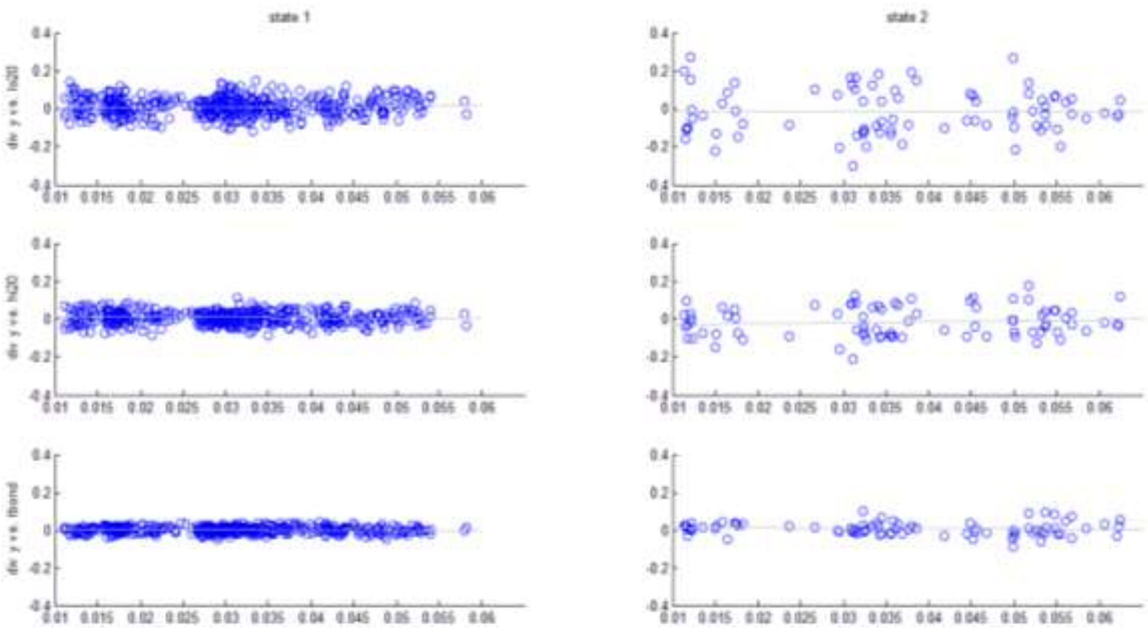


Figure 38 scatter of the dividend yield vs. small, large and bonds returns



My results are also similar to what discovered by Henkel et al. (2011), they find predictability to be stronger in recessions; I found out that 9 out of 16 state 2 VAR(1) coefficients are larger than the state 1 counterparts in absolute terms, and that 7 out of 16 state 2 VAR(1) coefficients, against only 4 out of 16 in state 1, are significantly different from zero. The estimated state dependent VAR(1) matrix shows significant time variations in the ability of the dividend yield to predict future stock returns and in the prediction itself, it can be seen that higher dividend yields forecast higher stock returns in state 1, but negative ones in state 2. The autoregressive coefficients suggest significant predictive power of bonds and dividend yield in both states 1 and 2, while the autoregressive coefficients of large stock returns fail to be significant in both states and the autoregressive coefficients of small stock returns indicate significant predictability only in state 1. Lagged bonds returns have the strongest predictive power in state 1, while in state 2, characterized by less predictability of returns, there is statistical evidence that only large stocks and bonds lagged returns affect respectively small stocks and bonds returns.

The different properties of the dividend yield across the two state affect the conditional distribution of the other asset classes returns, even though in both states the estimates suggests a low predictability on the dividend yield, in fact the regression coefficients of any expected return time series, except for the div_y on

its own lag, are not statistically significant. The estimate of the regime dependent VAR(1) matrix suggests that a higher dividend yield in state 1 forecasts higher lo20 and hi20 returns, and lower tbond returns while in state 2 it forecasts lower lo20 and hi20 returns, and higher tbond returns. The dividend yield is highly persistent in fact its autoregressive coefficient estimate is 0.9875 in state 1 and 1.0234 in state 2. Given the high persistence in the dividend yield time series, a single lag is required for the model.

3.8.3 A comparative analysis between the estimated regimes and the NBER USA recession indicator

In the following subparagraph a comparative analysis between the occurrence of recession periods estimated by my model and the NBER recession periods is illustrated. The analysis has been conducted using data provided by the National Bureau of Economic Research in addition to the data provided by the estimation of my model. The NBER based Recession Indicators for the United States from the Peak through the Period preceding the Trough is an interpretation of US Business Cycle Expansions and Contractions data provided by The National Bureau of Economic Research (NBER) and realized by the Federal Reserve Bank of St. Louis, the indicator essentially classifies each month either as a recession (value equal to 1) or an expansion period (value equal to 0). I also used another time series provided by the Federal Reserve Bank of St. Louis namely The Smoothed U.S. Recession Probabilities that is obtained from a dynamic-factor Markov switching model applied to four monthly coincident variables: non-farm payroll employment, the index of industrial production, real personal income excluding transfer payments, and real manufacturing and trade sales. This model is developed by Piger, Jeremy Max, Chauvet and Marcelle[13]. This time series essentially indicates for each month the probability that the United States economy is experiencing a recession. Starting from these two time series provide by the Federal Reserve Bank of St. Louis I constructed two time series that represent the NBER expansion indicator and a NBER smoothed U.S expansion probabilities; the

[13] This model was originally developed in Chauvet, M., "An Economic Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching," International Economic Review, 1998, 39, 969-996.

former one is the complement to 1 of the NBER recession indicator assuming value 1 during expansion period and 0 during recession periods, the latter one is the complement to 1 of the NBER smoother U.S. recession probabilities.

Figure 39 estimated states from the multivariate markow switching model

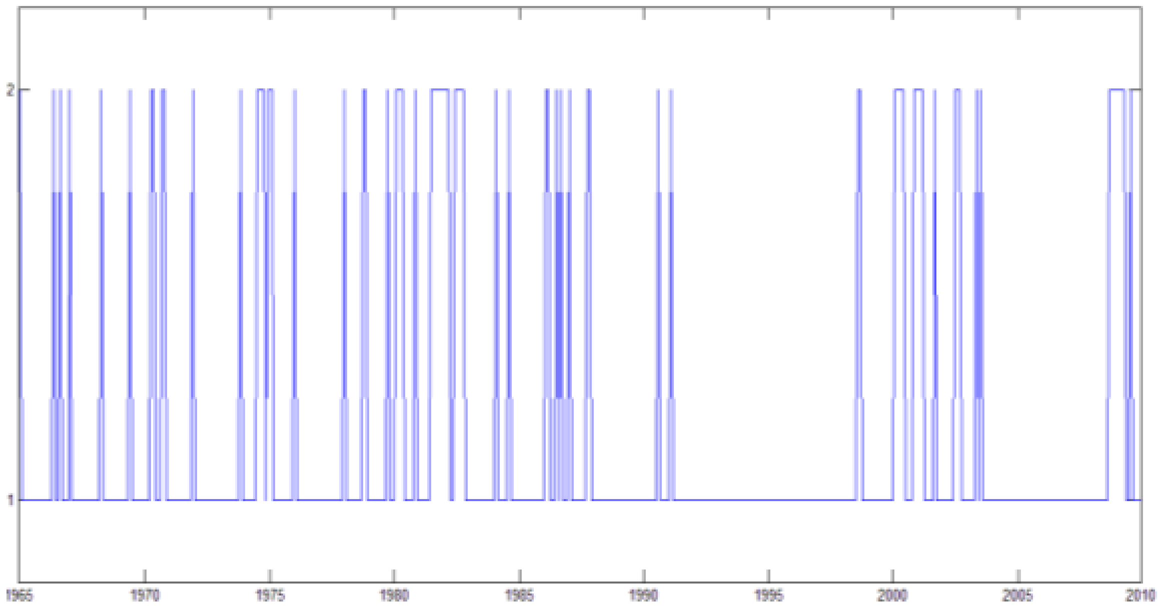


Figure 40 NBER based Recession Indicators for the United States from the Peak through the Period preceding the Trough (USRECP)

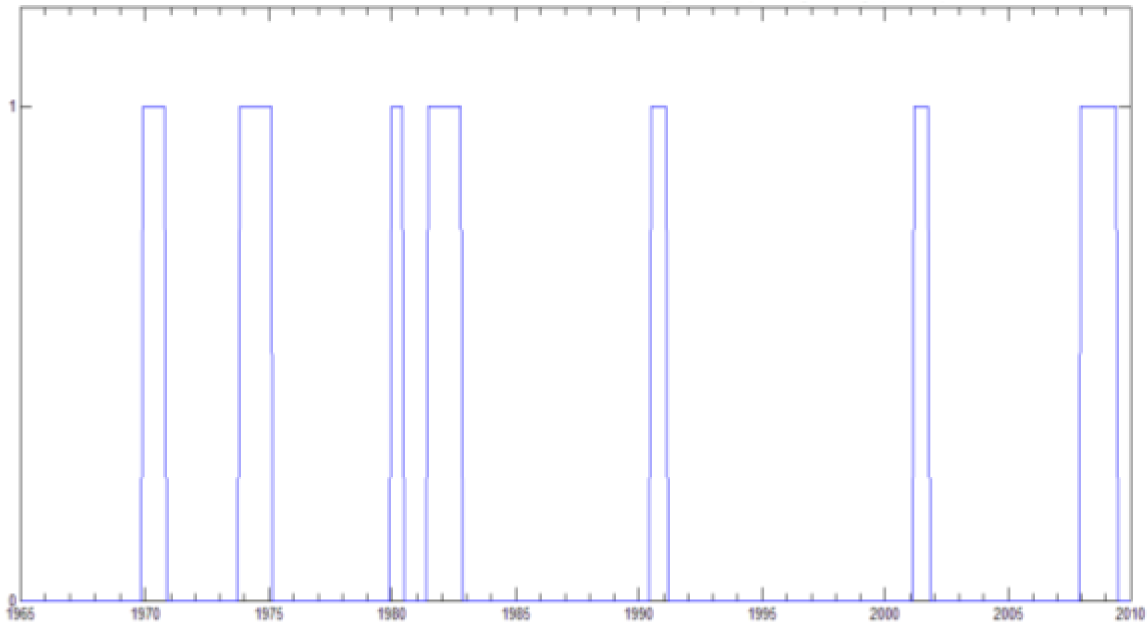


Figure 41 estimated smoothed states probabilities

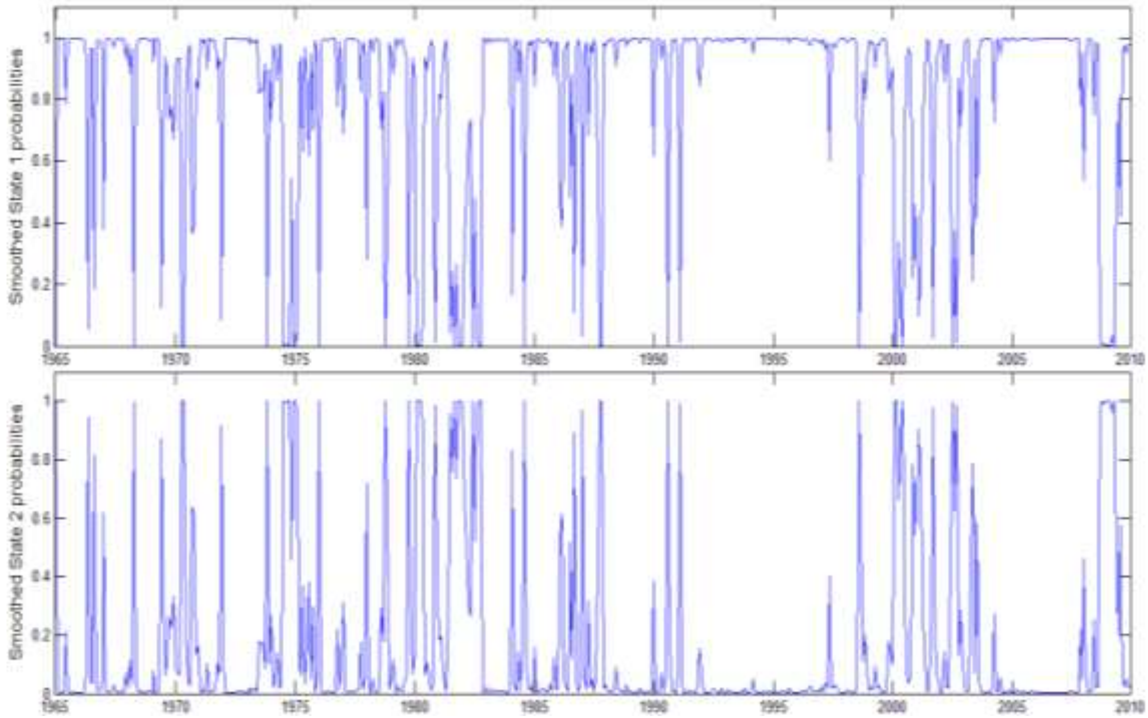


Figure 42 NBER Smoothed U.S. Recession Probabilities

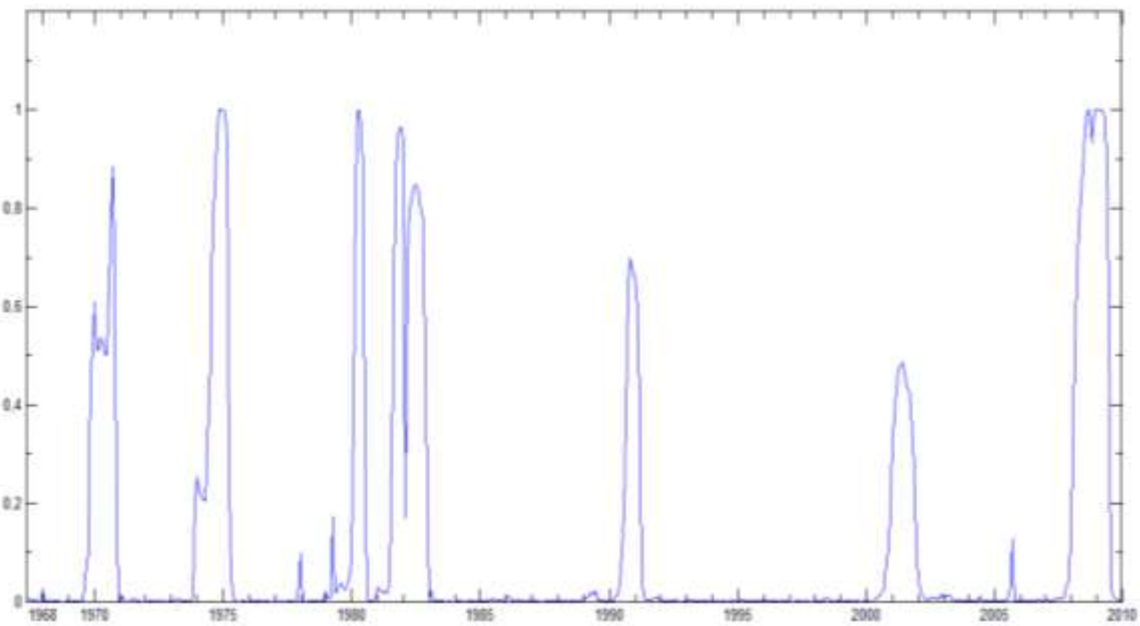


Figure 43 asset classes returns, dividend yield, dividend yield changes and state 2 probabilities

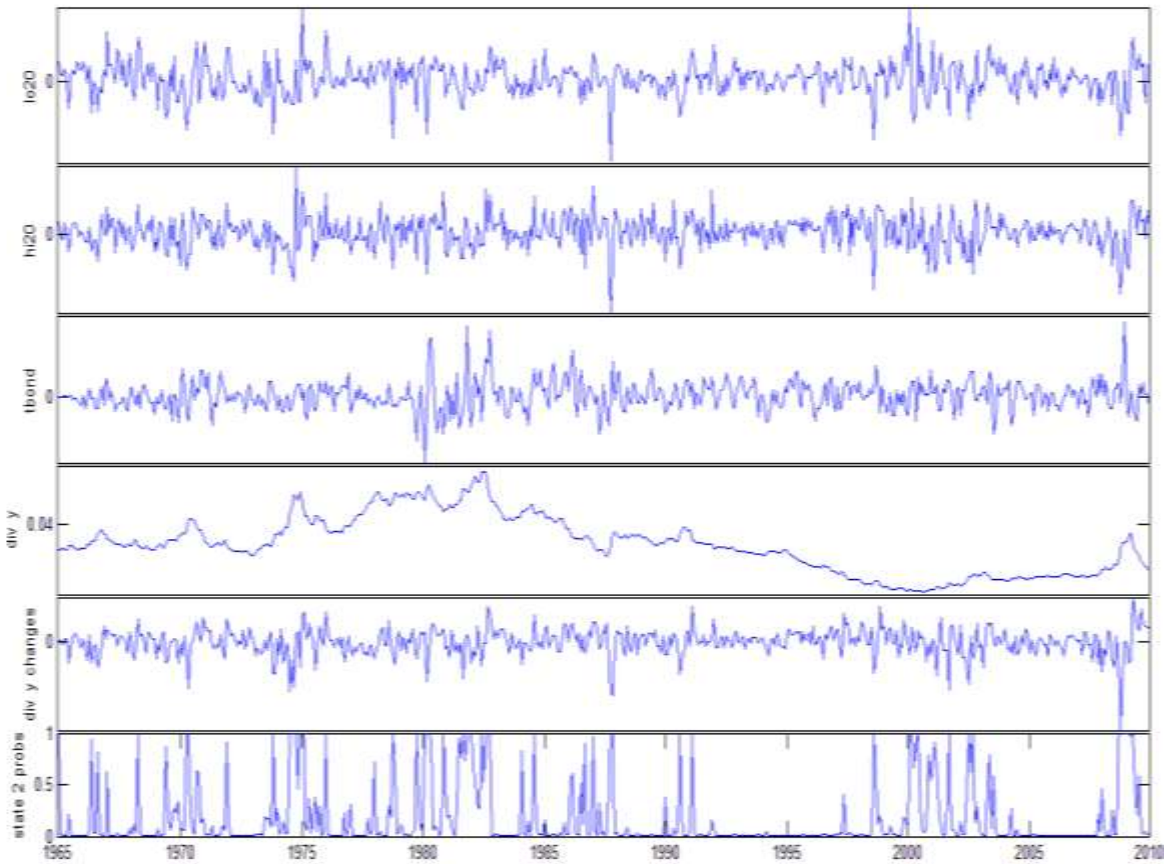
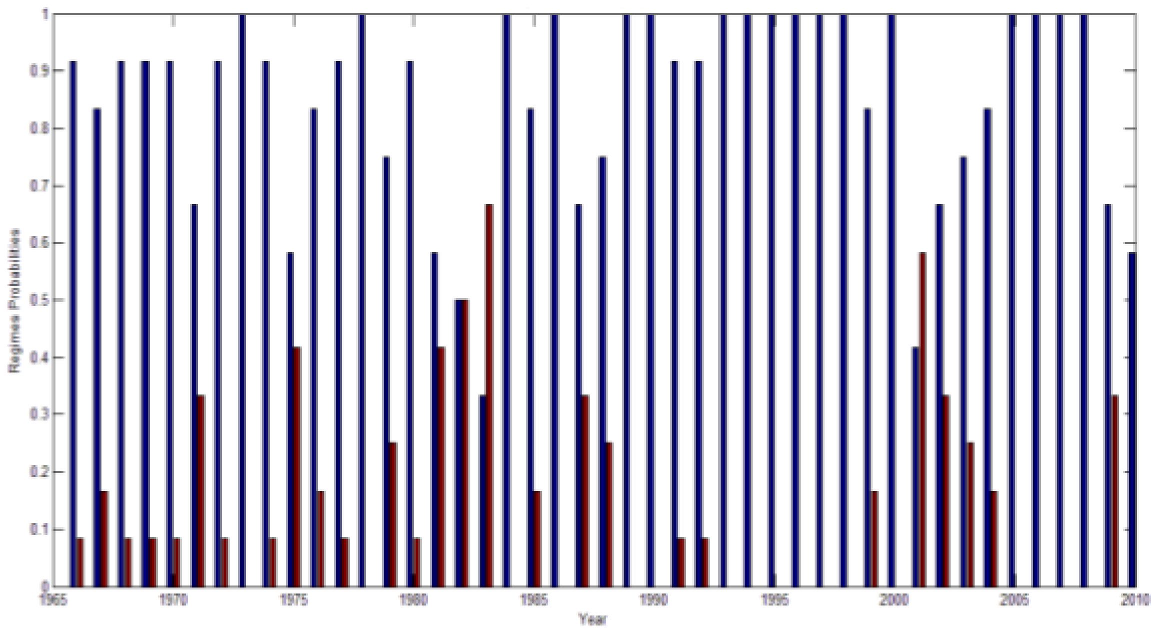


Figure 44 states frequency occurrence by year



The figure above represents how often the estimated model stays in a regime on average in a year. Each annual value relative to a certain state is the ratio

between the number of months in which the model has been in a that state and the number of months in a year, i.e. 12.

Table 18 frequencies of common NBER and MODEL's recession and expansion periods

	Model recession periods	Model expansion periods
NBER recession periods	0.518519	0.089325
NBER expansion periods	0.481481	0.910675

As evidenced in Table 18, I found that approximately 52% of the periods classified as recession (state 2) by my model occur in an NBER recession period, conversely approximately 91% of the periods classified as expansion (state 1) occur in an NBER expansion period.

Table 19 regression of estimated state 1 smoothed probabilities on NBER recession and expansion indicator

	NBER recession indicator	NBER expansion indicator
smoothed state 1 probabilities	-0.484702	0.484702
smoothed state 2 probabilities	0.484702	-0.484702

As shown in Table 19 correlation between estimated smoothed state probabilities and NBER recession dates are -0.484702 for state 1 and 0.484702 for state 2; as observed by Guidolin and Timmermann (2005) since the state probabilities sum to one, by construction, if some correlations are positive, others must be negative. This evidences bring other support to the fundamental idea that state 1 occurs around official recession periods indicated by the NBER recession indicator.

Figure 45 regression of estimated state 1 smoothed probabilities on NBER recession indicator

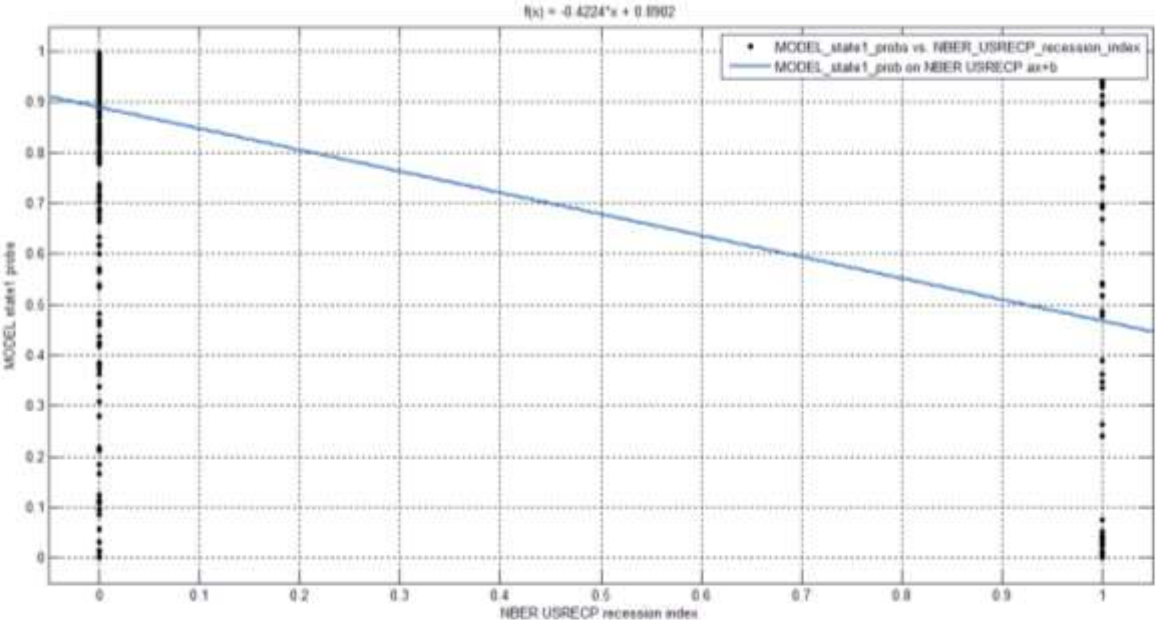
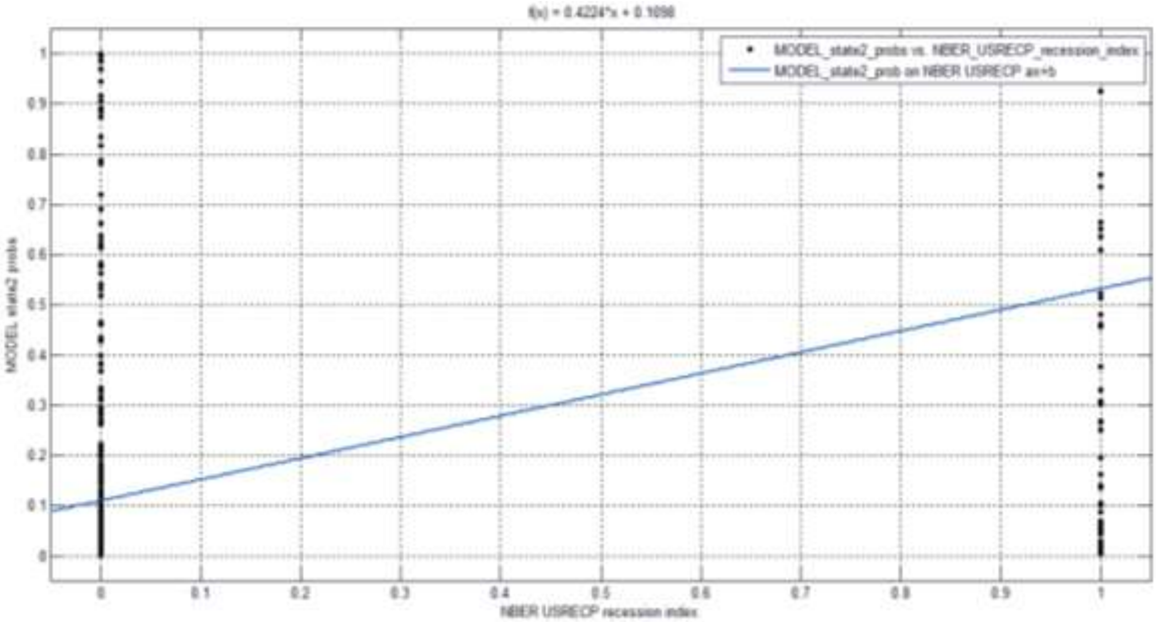


Figure 46 regression of estimated state 2 smoothed probabilities on NBER recession indicator



The regression analysis of estimated state probabilities on the NBER recession indicator, illustrated in Figure 45 and Figure 46, leads to the conclusion that in the NBER recession periods the estimated state 1 smooth probabilities assumes large values while state 2 smooth probabilities assumes small values; the opposite situation occurs in the NBER expansion periods. In fact the regression analysis of

states probabilities on the NBER indicator shows a negative coefficient for state 1 and a positive coefficient for state 2.

Figure 47 regression of estimated 1 month lagged state 1 smoothed probabilities on NBER recession indicator

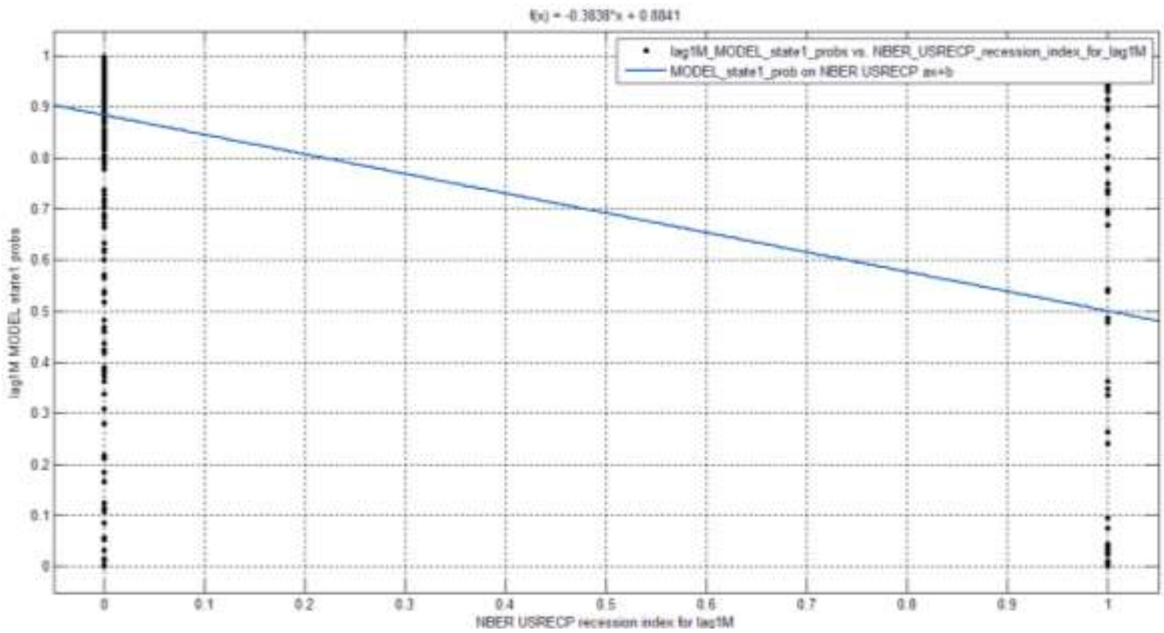


Figure 48 regression of estimated 1 month lagged state 2 smoothed probabilities on NBER recession indicator

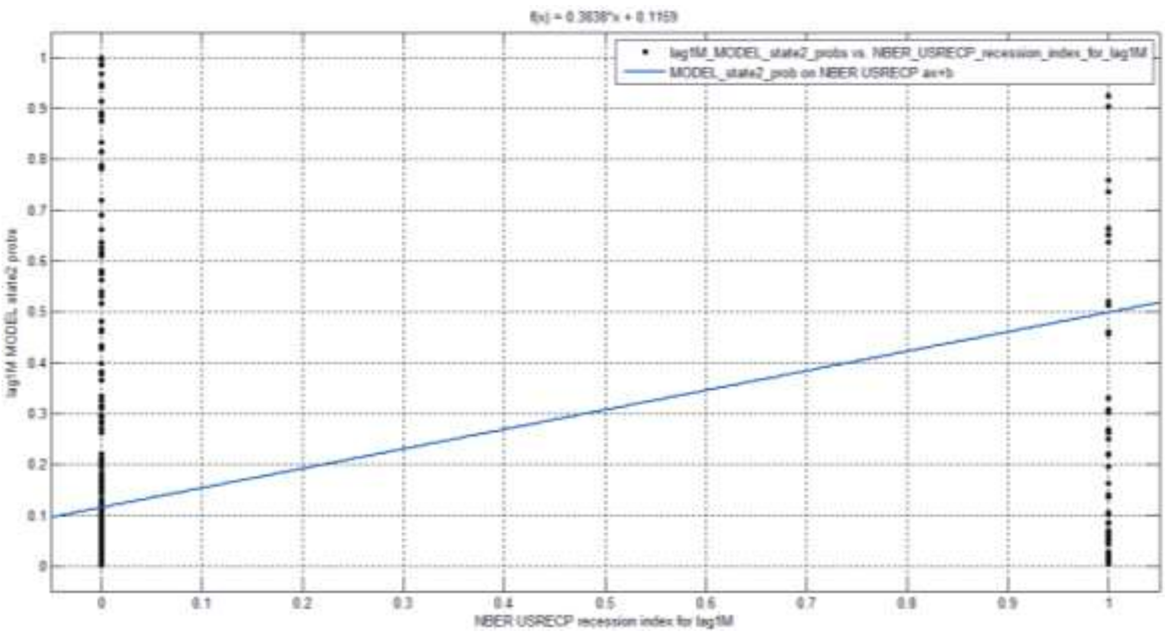


Figure 49 regression of estimated 3 months lagged state 1 smoothed probabilities on NBER recession indicator

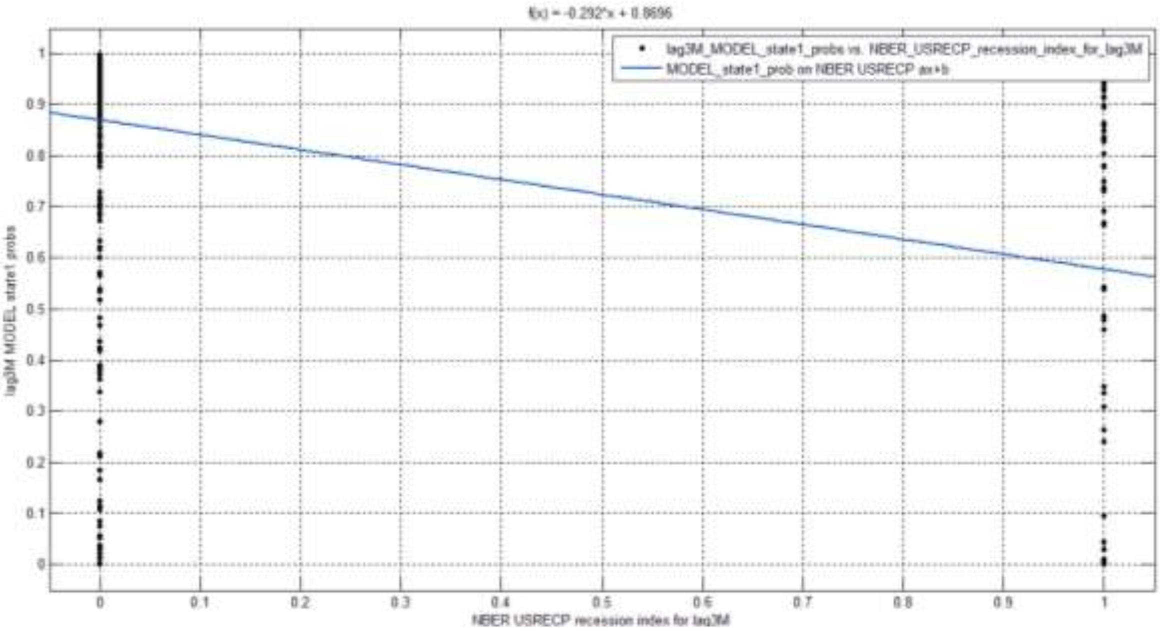


Figure 50 regression of estimated 3 months lagged state 2 smoothed probabilities on NBER recession indicator

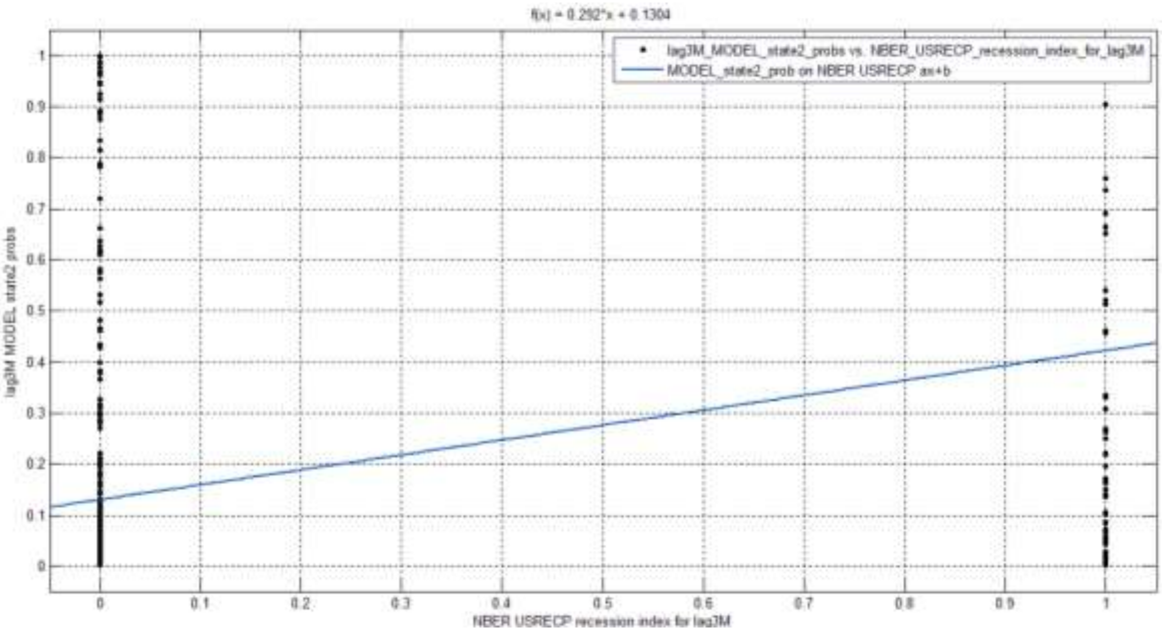


Figure 51 regression of estimated 6 months lagged state 1 smoothed probabilities on NBER recession indicator

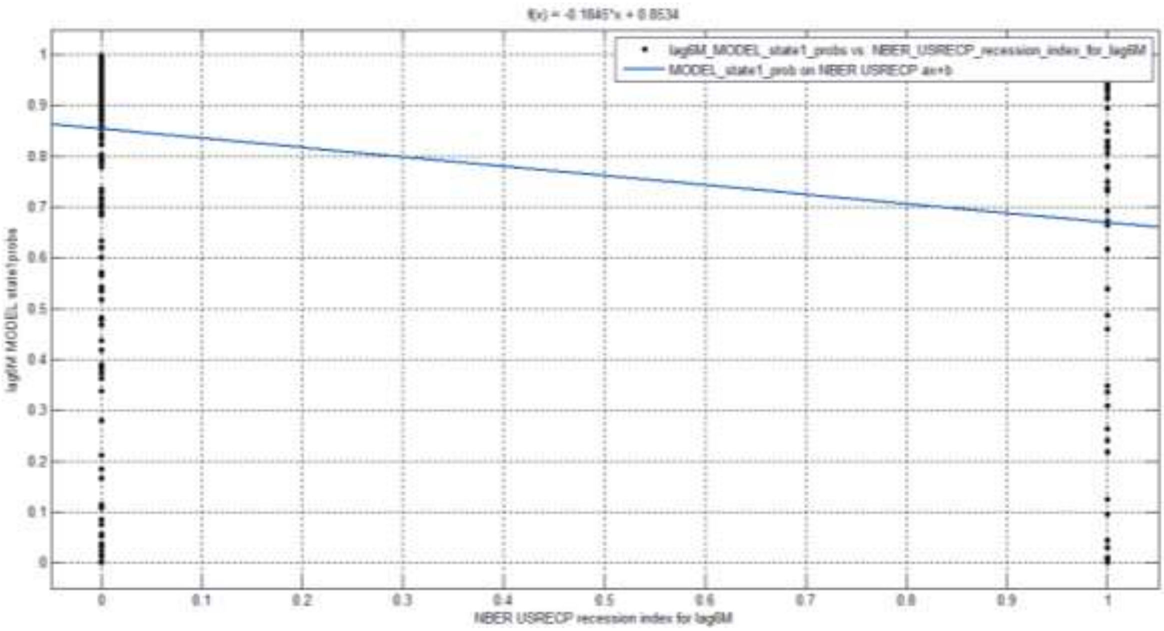
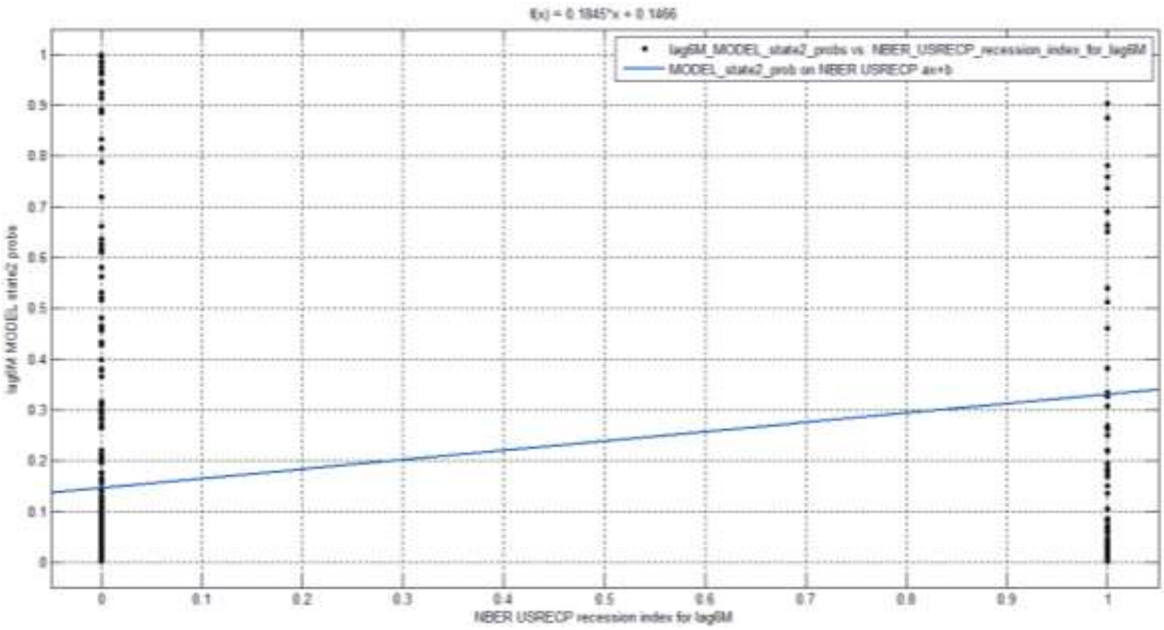


Figure 52 regression of estimated 6 months lagged state 2 smoothed probabilities on NBER recession indicator



As suggested by Guidolin and Timmermann (2005) it may be argued that the state probabilities estimated from financial returns should lead economic recession months, for this reason I also conducted a regression analysis of state 1 and state 2 probabilities lagged 1, 3 and 6 months on the NBER recession indicator. The results are shown in Figure 47, 48, 49, 50, 51 and 52. It can be clearly concluded that both the state 1 and state 2 regression coefficients show a decrease in

absolute terms when the lag become larger, therefore the larger the lag in the state probabilities time series the weaker the positive relation between the lagged recession probabilities and the presence of a recession period in the NBER recession indicator, this means that there is no evidence of an improved capacity of the estimated lagged recession probabilities, compared to the not lagged one, to forecast the presence of a recession in the NBER recession indicator.

Figure 53 regression of estimated state 1 (expansion) smoothed probabilities on NBER expansion probabilities

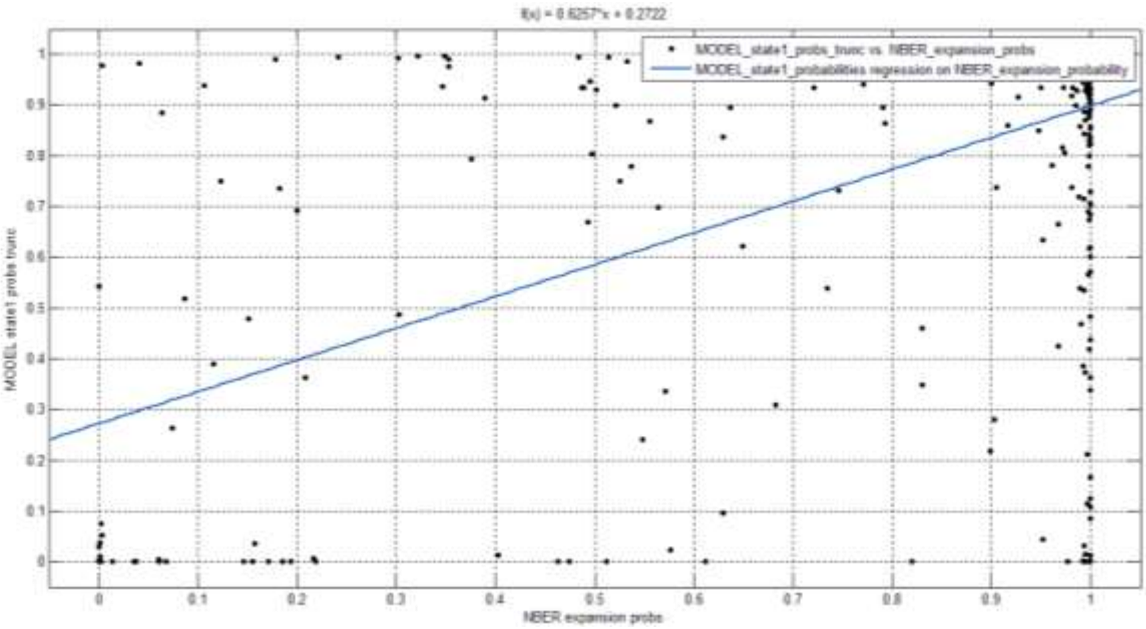


Figure 54 regression of estimated state 1 (expansion) smoothed probabilities on NBER recession probabilities

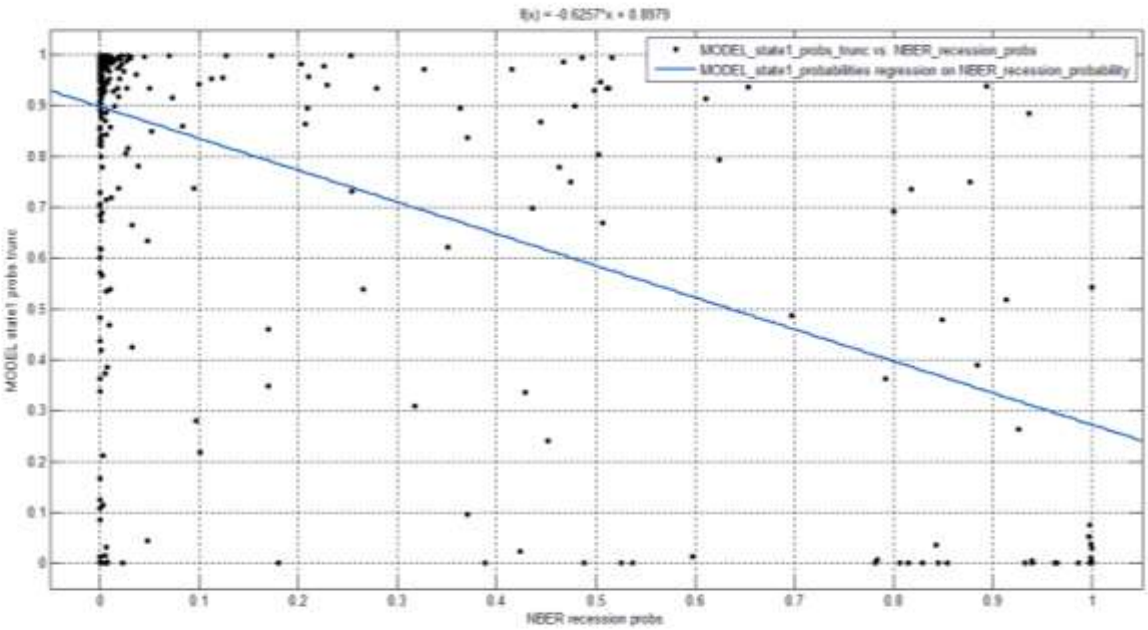


Figure 55 regression of estimated state 2 (recession) smoothed probabilities on NBER expansion probabilities

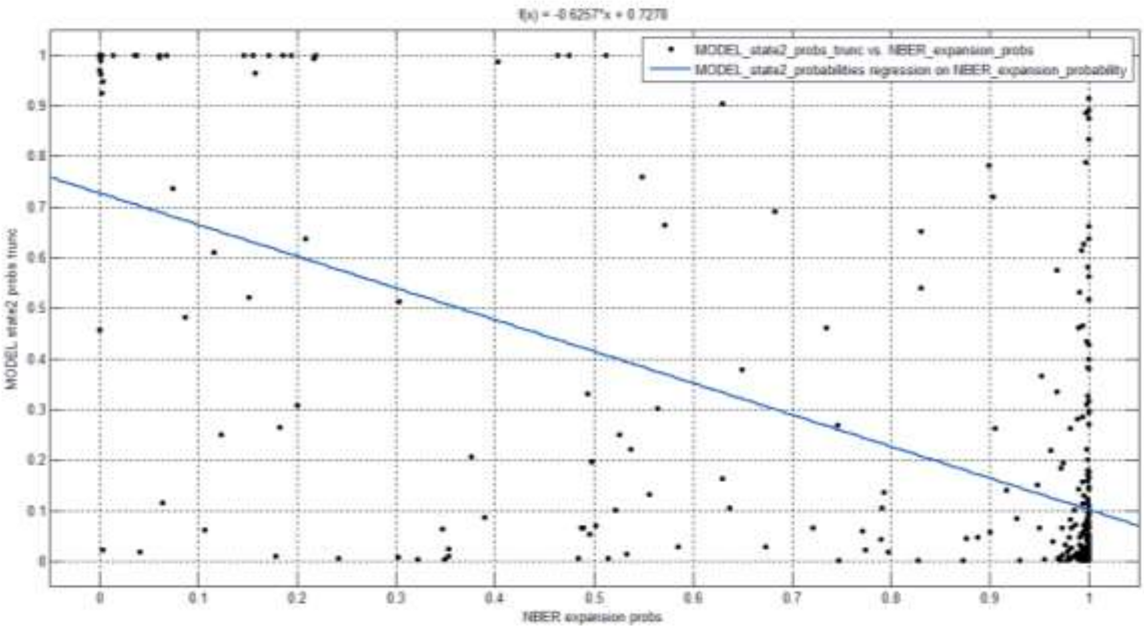
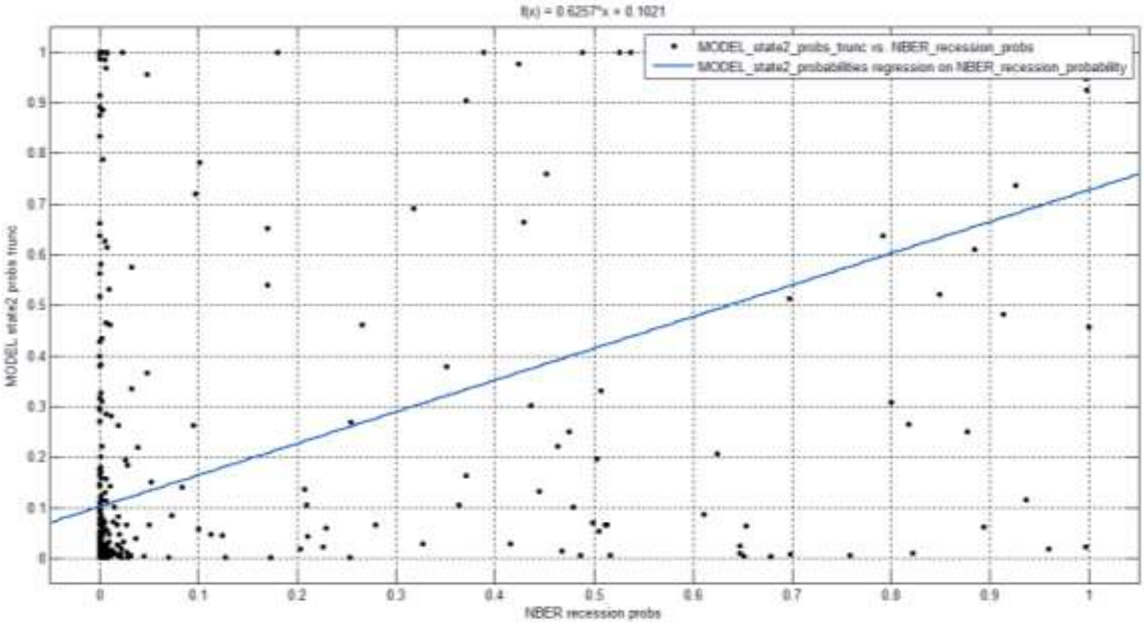


Figure 56 regression of estimated state 2 (recession) smoothed probabilities on NBER recession probabilities



I conducted an additional regression analysis of both the state 1 and state 2 probabilities on the NBER recession and NBER expansion probabilities, the results are shown in Figure 53, 54, 55 and 56. The NBER recession probabilities is represented by the Smoothed U.S. Recession Probabilities supplied by the American National Bureau of Economic Research with monthly frequency from June 1967. This analysis further supports the fundamental idea that the estimated state 2 represents recession periods while state 1 represents expansion period. The regression analysis of the estimated state 1 probabilities on the NBER expansion probabilities and NBER recession probabilities respectively shows a large and positive and a large and negative coefficient. The antithetical situation occurs with the regression analysis of the estimated state 2 probabilities.

Table 20 correlations between estimated smoothed state probabilities and NBER recession and expansion probabilities

	NBER recession probabilities	NBER expansion probabilities
state 1 smoothed probabilities	-0.533960	0.533960
state 2 smoothed probabilities	0.533960	-0.533960

The same conclusion is also supported by the correlations between the estimated smoothed state probabilities and the NBER recession and expansion probabilities shown in Table 20, it can be seen that there is evidence of a positive relation between the state 2 (bear market, recession) estimated smoothed probabilities and the NBER recession probabilities and a negative relation between the state 1 (bull market, expansion) estimated smoothed probabilities and the NBER recession probabilities. All the evidences presented so far suggest that the regime switching estimate of my model appears to be related to the underlying economic fundamentals and business cycle to some extent. As suggested by Jun Tu (2010) a possible explanation of the divergence between the regime switching and the underlying business cycle may come from the fact that stock markets also react to sectoral or shorter-lived contractions in the economy that are not designated as recessions by NBER.

3.8.4 A dynamic correlation analysis of asset classes returns

In this subparagraph I am going to illustrate the findings relative to the dynamic correlation analysis I have realized. I compute and plot the dynamics of the conditional correlations implied by the model using only real time information (i.e., using filtered and not smoothed probabilities). In computing dynamic correlations it has been necessary to adopt some adjustment as suggested by Guidolin in "Modelling, Estimating and Forecasting Financial Data under Regime (Markov) Switching", in order to take into account the effects of both variances and covariances of the joint presence of switches in expected excess returns. The covariance matrices of the regime switching restricted MSVAR(2,1) model I have estimated are characterized by the fact that the covariances in both regimes are restricted to be zero; as a consequence the only source of correlation between the asset classes time series in the system is due to the presence of a unique Markov switching dynamic, in common for all the asset returns, which drives the moments of three asset classes time series. The additional source of correlation due to the presence of a common Markov switching dynamic comes from the fact that the Markov state moves the means in the same direction, except or the large

stocks, and at the same time, which makes the standard correlation an imperfect measure of comovements.

The dynamic correlation formula is here illustrated:

(38)

$$\rho_t^{a,b} = \frac{\sigma_{ab,1}\xi_{1t}+\sigma_{ab,2}\xi_{2t}+\xi_{1t}\xi_{2t}(\mathrm{E}(a_{t+1}|\boldsymbol{y_t},\Im_t;\boldsymbol{\theta})-\mathrm{E}(a))(\mathrm{E}(b_{t+1}|\boldsymbol{y_t},\Im_t;\boldsymbol{\theta})-\mathrm{E}(b))}{\left(\left(\sigma_{a,1}^2\xi_{1t}+\sigma_{a,2}^2\xi_{2t}+\xi_{1t}\xi_{2t}\left(\mathrm{E}(a_{t+1}|\boldsymbol{y_t},\Im_t;\boldsymbol{\theta})-\mathrm{E}(a)\right)^2\right)\left(\sigma_{b,1}^2\xi_{1t}+\sigma_{b,2}^2\xi_{2t}+\xi_{1t}\xi_{2t}\left(\mathrm{E}(b_{t+1}|\boldsymbol{y_t},\Im_t;\boldsymbol{\theta})-\mathrm{E}(b)\right)^2\right)\right)}$$

where a and b are two asset classes time series, ξ_{it} is the filter probability of regime *i* at time *t*, $\sigma_{ab,i}$ is covariance between a and b in state *i*, $\sigma_{a,i}^2$ is the variance of a in state *i*, $\mathrm{E}(a) = \mathrm{E}(a_{t+1}|\boldsymbol{\theta}) = \xi_{1\infty}\bar{\mu}_{a,S_t=1} + (1-\xi_{1\infty})\bar{\mu}_{a,S_t=2}$ is the unconditional mean of a (formula (32)), while $\mathrm{E}(a_{t+1}|\boldsymbol{y_t},\Im_t;\boldsymbol{\theta})$ is the conditional mean of a based on the information available at time *t*. The information available at time *t* consists of $\boldsymbol{y_t}$, a vector that contains the values of a,b and c at time *t*, and the filtered probabilities ξ_{1t} and ξ_{2t}. The conditional mean formula is here illustrated:

(39) $\mathrm{E}(a_{t+1}|\boldsymbol{y_t},\Im_t;\boldsymbol{\theta}) = (\xi_{1t}\mathrm{p}_{11} + \xi_{2t}\mathrm{p}_{12})\mathrm{E}(a_{t+1}|\mathrm{S}_t = 1,\boldsymbol{y_t},\Im_t;\boldsymbol{\theta}) + (\xi_{1t}\mathrm{p}_{21} + \xi_{2t}\mathrm{p}_{22})\mathrm{E}(a_{t+1}|\mathrm{S}_t = 2,\boldsymbol{y_t},\Im_t;\boldsymbol{\theta})$

where (assuming that a is modeled by a MSVAR(2,1) and that there are another asset, b, and a predictor c)

(40) $\mathrm{E}(a_{t+1}|\mathrm{S}_t = j,\boldsymbol{y_t},\Im_t;\boldsymbol{\theta}) = \mu_{a\mathrm{S}_t=j} + \phi_{\mathrm{S}_t=j}^{a,a}a_t + \phi_{\mathrm{S}_t=j}^{a,b}b_t + \phi_{\mathrm{S}_t=j}^{a,c}c_t$

The conditional mean term also represents, as illustrated at the beginning of the paragraph, the mean of the normal distribution

(41) $f(R_t|\Im_{t-1};\boldsymbol{\theta}) = \sum_{j=1}^{2}\sum_{i=1}^{2}\mathrm{p}_{ji}\,\xi_{it-1}\eta_{jt}$

which represents the conditional density of the *t*-th observation. The above correlation coefficient is calculated for each period, as a result a time series of length 540 is obtained for the dynamic correlation between any two asset classes time series.

The numerator in the ratio calculation is the filtered covariance at time *t* adjusted to take into account the effects of regime switches while the terms in the denominator are the filtered standard deviation at time *t* adjusted to take into account the effects of regime switches. Similarly to the conditional mean, the conditional standard deviation is here illustrated:

(42) $$Var(\boldsymbol{y}_{\mathbf{t+1}}|\Im_{\mathrm{t}};\boldsymbol{\theta}) = \sigma_{a,t+1|\Im_{\mathrm{t}};\boldsymbol{\theta}} = (\xi_{1\mathrm{t}}\mathrm{p}_{11} + \xi_{2\mathrm{t}}\mathrm{p}_{12})\sigma_{a,\mathrm{S}_{\mathrm{t}}=1} + (\xi_{1\mathrm{t}}\mathrm{p}_{21} + \xi_{2\mathrm{t}}\mathrm{p}_{22})\sigma_{a,\mathrm{S}_{\mathrm{t}}=2}$$

Figure 57 model implied dynamic correlation between small stocks and bonds

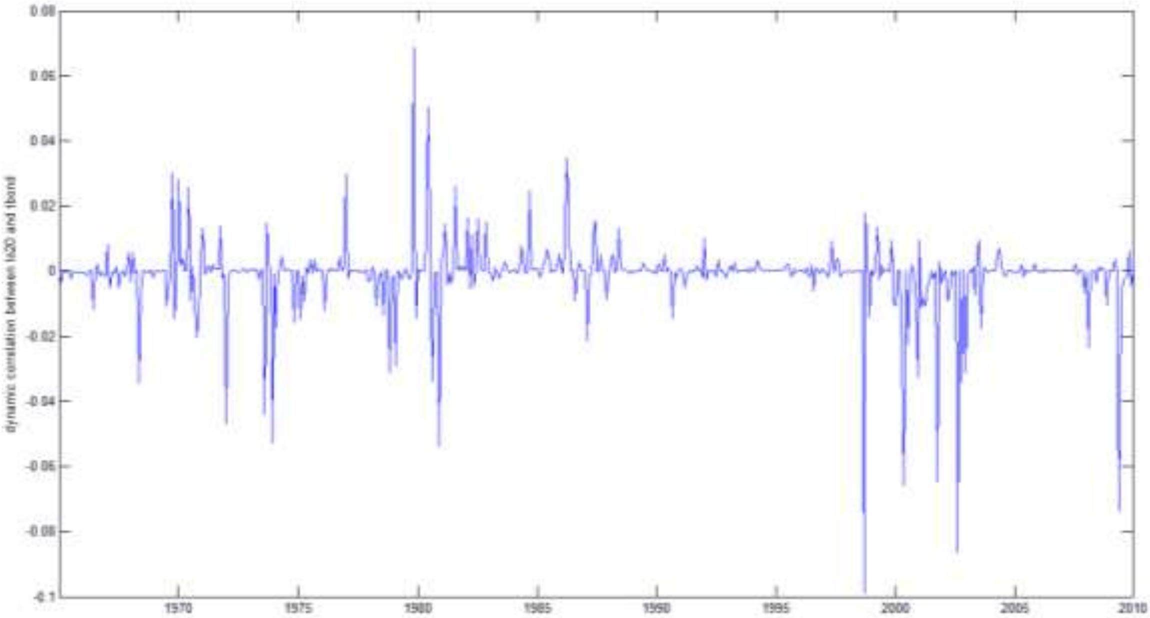


Figure 58 model implied dynamic correlation between small stocks and large stocks

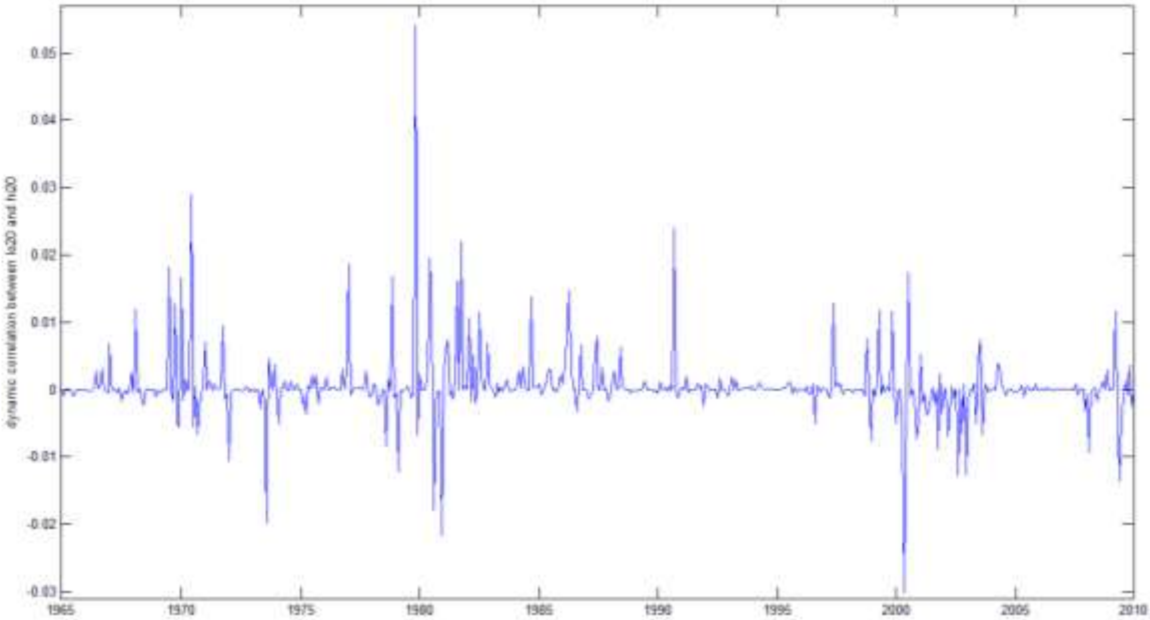


Figure 59 model implied dynamic correlation between large stocks and bonds

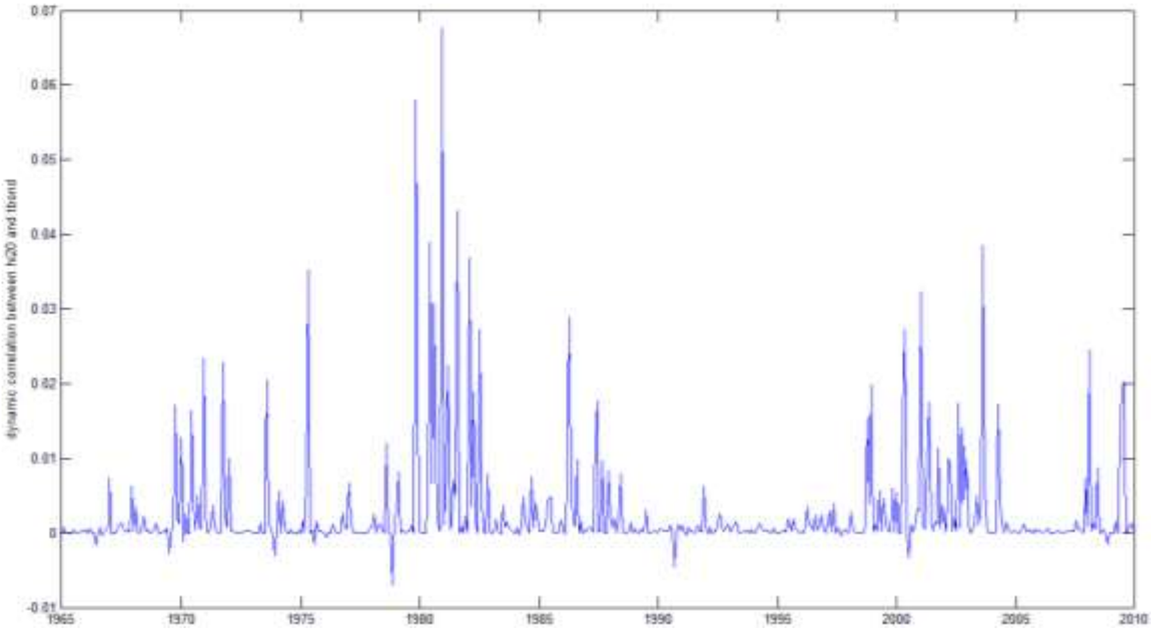


Figure 60 dynamic correlation between asset classes returns and dividend yield

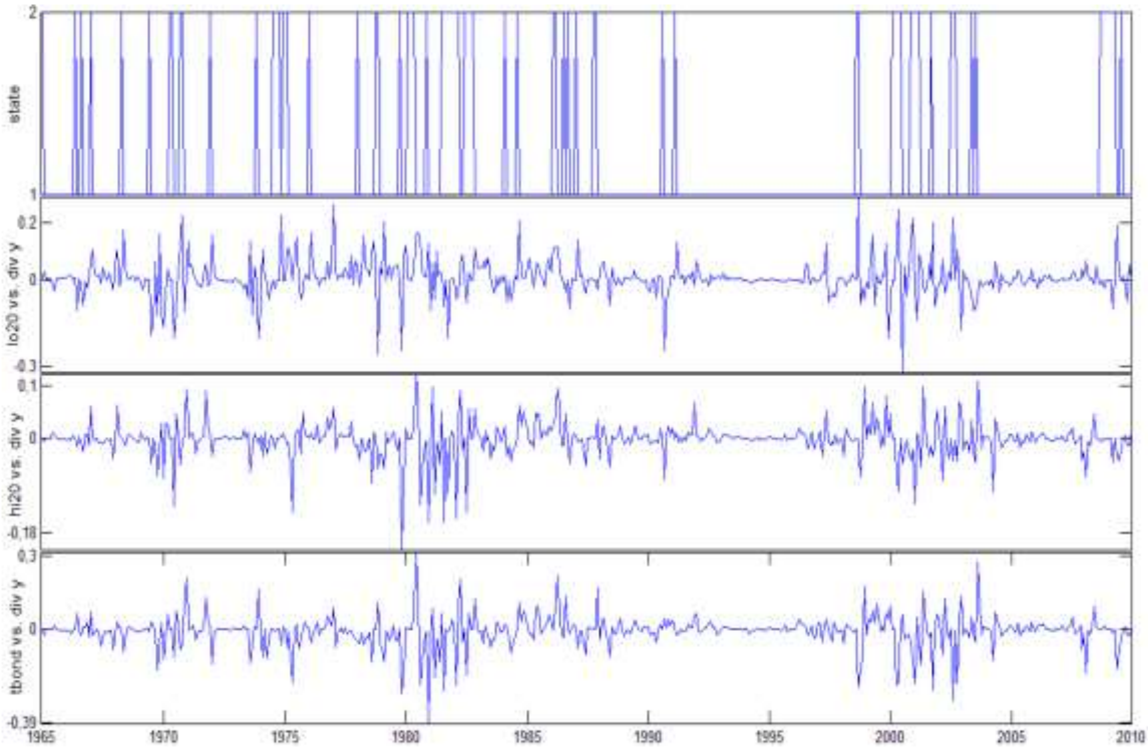


Table 21 unconditional and state conditional average dynamic correlation coefficients

STATE	lo20 vs. tbond	lo20 vs. hi20	hi20 vs. tbond	lo20 vs. div_y	hi20 vs. div_y	tbond vs. div_y
state 1	-0.0007	0.0004	0.0024	0.0070	-0.0013	-0.0062
state 2	-0.0040	0.0006	0.0035	0.0109	-0.0109	-0.0234
unconditional	-0.0012	0.0005	0.0026	0.0076	-0.0027	-0.0088

Table 21 shows the unconditional and state conditional average dynamic correlation coefficients. The first two records of the table contain the state conditional average dynamic correlation coefficients, each of them calculated as the average values of the corresponding dynamic correlation time series, conditionally on the state value contained in the first column ; the third record of the table shows the unconditional average dynamic correlation coefficient calculated as the average value of the corresponding whole dynamic correlation time series. Overall, it can be seen from Table 21, and more clearly from Figure 93, that the average dynamic correlation coefficients are, in absolute term, substantially larger in state 2 than in state 1, this finding confirms the already discussed tendency of the return to comove greatly in the recession regime (state

2). For each pair of time series the unconditional average dynamic correlation coefficient assumes a value between the two corresponding state conditional average values. From a comparison between Table 21 and Table 16 it emerges that 5 out of 6 correlation coefficients between the dividend yield and an asset class share the same arithmetic sign, thus the two different methodologies used to produce the results in the two tables bring to the same conclusion; the exception is represented by the correlation coefficient between the large stocks and the dividend yield in state 2. As it can be seen from Table 11, the dividend yield and the small stocks state conditional means are larger in state 2 than in state 1, while small stocks and bonds state conditional means are larger in state 2 than in state 1. In this context, during a recession period (state 2) the dividend yield is above its unconditional mean as well as the small stock returns, thus a positive comovement, as illustrated in Table 21, occurs; at the same time the large stocks perform worse in state 2 than in state 1, thus the dividend yield and the large stock returns move to different direction; lastly the comovement in the opposite direction of bonds and dividend yield can be explained by the significant positive comovement between large stocks and bonds that make bonds returns grow, together with large stock returns, when the economy is experiencing an expansion period and the dividend yield is low. The dynamic just explained is partly originated by the time-variations in the dividend yield, which may induce a large hedging demand.

3.8.5 Regime shifts in the asset classes returns means and volatilities

The following charts show the asset classes returns and their relative time varying conditional mean and time varying conditional standard deviations estimated by the model.

Figure 61 asset classes returns and estimated time varying conditional means

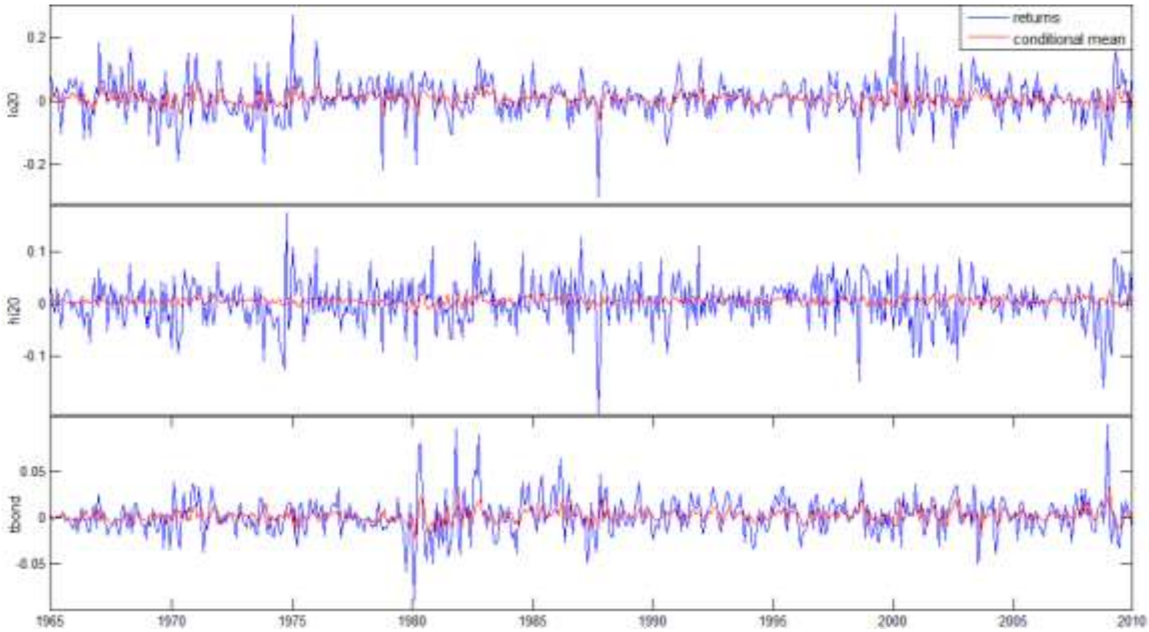


Figure 62 estimated time varying conditional standard deviations

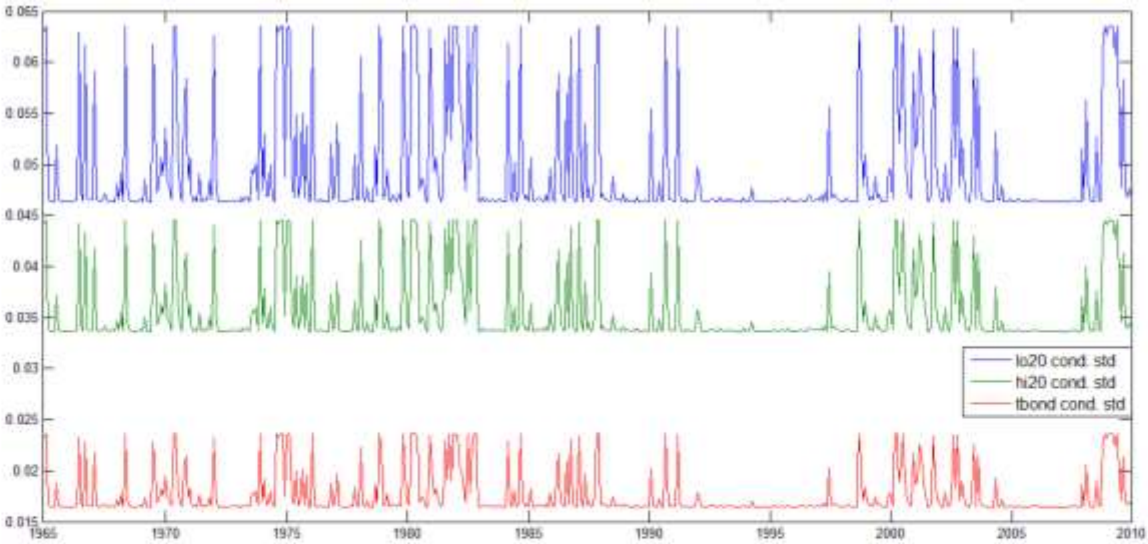
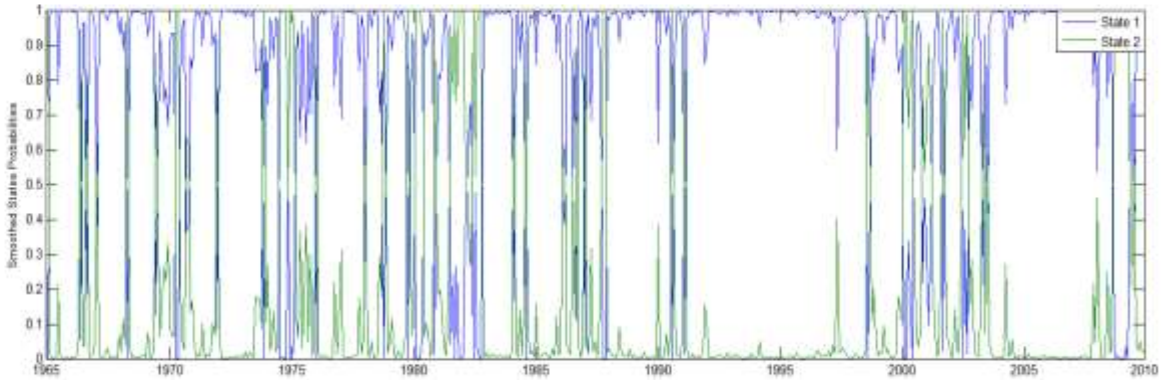


Figure 63 estimated smoothed state probabilities



The two plots above show that the three asset classes expected excess returns, their conditional means and their conditional standard deviations are to some extent synchronized across time series, reflecting the shapes of the evolution of smoothed probabilities.

3.9 A single regime VAR(1) model estimates

In order to evaluate the potential benefit of a multiple regimes model against a more simple single regime model, a first order vector autoregressive model VAR(1) has been estimated. A comparison of the two model performances, both in the in-sample and out-of-sample framework, has been conducted.

The Markov Switching Vector Autoregressive model MSVAR is a general class of models which nests the standard VAR model but additionally accounts for nonlinear regime shifts. As a consequence the VAR(1) model is a special case of the MSVAR(2,1) in which the intercepts terms are, the VAR(1) autoregressive coefficients and the covariance matrix are not regime-dependent, and then the model is homoskedastic. This model configuration shows a static and simultaneous contagious effect through the off-diagonal elements of the variance and covariance matrix, a static and linear one through the VAR coefficients. The VAR(1) model is based on the following expression:

(43) $\boldsymbol{y}_{\mathbf{t}} = \boldsymbol{\mu} + \boldsymbol{A}_{\mathbf{1}}\boldsymbol{y}_{\mathbf{t-1}} + \boldsymbol{\varepsilon}_{\mathbf{t}}$

where $\boldsymbol{y}_{\mathbf{t}} = (\mathbf{r}_{\mathbf{t}}' \ \mathbf{z}_{\mathbf{t}}')'$ is a *(3 + 1) x* 1 vector that contains lo20, hi20 and tbond returns and an external predictor $\mathbf{z}_{\mathbf{t}}$ which is represented by the dividend yield

div_y; $\boldsymbol{\mu} = [\mu_1 \ldots \mu_4]'$ is the intercept vector for $\boldsymbol{y}_{\mathbf{t}}$, $\boldsymbol{A}_{\mathbf{1}}$is a *(3 + 1) x (3 + 1)* matrix of autoregressive and regression coefficients while $\boldsymbol{\varepsilon}_{\mathbf{t}} = [\varepsilon_{1t} \ldots \varepsilon_{4t}]' \sim N(\mathbf{0}, \boldsymbol{\Sigma})$.

The model has been estimated using the in-sample data from January 1965 to December 2010; Table 22 contains the parameter estimates.

Table 22 single regime VAR(1) parameter estimates

	VAR(1)			
	lo20	hi20	tbond	div_y
1. Intercept term	-0.0040 (0.0075)	-0.0026 (0.0052)	0.0005 (0.0023)	0.0007 (0.0002)
2. VAR(1) Matrix				
lo20	0.1665 (0.0600)	0.1234 (0.0905)	0.1428 (0.1376)	0.2910 (0.2244)
hi20	0.0434 (0.0412)	-0.0225 (0.0622)	0.2631 (0.0945)	0.1780 (0.1542)
tbond	-0.0509 (0.0181)	0.0096 (0.0272)	0.3235 (0.0414)	0.0280 (0.0675)
div_y	-0.0007 (0.0016)	-0.0148 (0.0024)	-0.0098 (0.0036)	0.9823 (0.0058)
3. Covariance Matrix				
lo20	0.003959			
hi20	0.001974	0.001869		
tbond	0.000167	0.000155	0.000358	
div_y	-0.000037	-0.000030	-0.000005	0.000003

The stability and invertibility of the model have been tested; given that all eigenvalues of the associated lag operators have modulus less than 1 it can be said that the model is stable and invertible.

CHAPTER 4 - ASSET ALLOCATION

4.1 Unconditional and state conditional asset classes returns distributions and efficient frontiers based on a MSVAR(2,1) model

The following charts show the asset classes returns and their relative time varying conditional means and time varying conditional standard deviations estimated by the model. The unconditional distributions are mixtures of the two state conditional normal distributions in which the weighting factors are equal to the corresponding state ergodic probabilities. As it can be seen the mixture of the two normals generates a distribution characterized by skewness and kurtosis. Given that the investor is not interested in investing in the dividend yield but only in the small stocks, large stocks and bonds, the expected unconditional mean for each *t* consists of the first three elements of the following vector:

(44)[14] $\mathrm{E}(\boldsymbol{y_{t+1}}|\boldsymbol{\theta}) = \xi_{1\infty}\overline{\boldsymbol{\mu_1}} + (1-\xi_{1\infty})\overline{\boldsymbol{\mu_2}}$

and the unconditional covariance matrix adjusted for the regime structure is equal to the *(3 x 3)* left-upper part of the following matrix:

(45)[15] $\mathrm{Var}(\boldsymbol{y_{t+1}}|\boldsymbol{\theta}) = \xi_{1\infty}\boldsymbol{\Sigma_1} + (1-\xi_{1\infty})\boldsymbol{\Sigma_2} + \xi_{1\infty}(1-\xi_{1\infty})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})^{\mathrm{T}}$

where

(46) $\overline{\boldsymbol{\mu_1}} = \mathrm{E}(\boldsymbol{y_t}|\mathrm{S_t} = 1; \boldsymbol{\theta}) = (\mathbf{I_4} - \boldsymbol{A_{1,1}})^{-1}\boldsymbol{\mu_1}$ and

(47) $\overline{\boldsymbol{\mu_2}} = \mathrm{E}(\boldsymbol{y_t}|\mathrm{S_t} = 2; \boldsymbol{\theta}) = (\mathbf{I_4} - \boldsymbol{A_{1,2}})^{\boldsymbol{-1}}\boldsymbol{\mu_2}$

represent the state conditional asset classes means while $\xi_{1\infty}$ and $\xi_{2\infty} = (1-\xi_{1\infty})$ are the state ergodic probabilities. The following figures plot for each asset class the probability density function corresponding to the mixture of two normals that

[14,15] Massimo Guidolin "Modelling, Estimating and Forecasting Financial Data under Regime (Markov) Switching" download at "http://didattica.unibocconi.it/mypage/dwload.php?nomefile=Lecture_7_-_Markov_Switching_Models20130520235704.pdf"

draws from state 1 density function with probability $\xi_{1\infty}$ and state 2 density function with probability $\xi_{2\infty}$.

Figure 64 small stocks state conditional and unconditional estimated density function

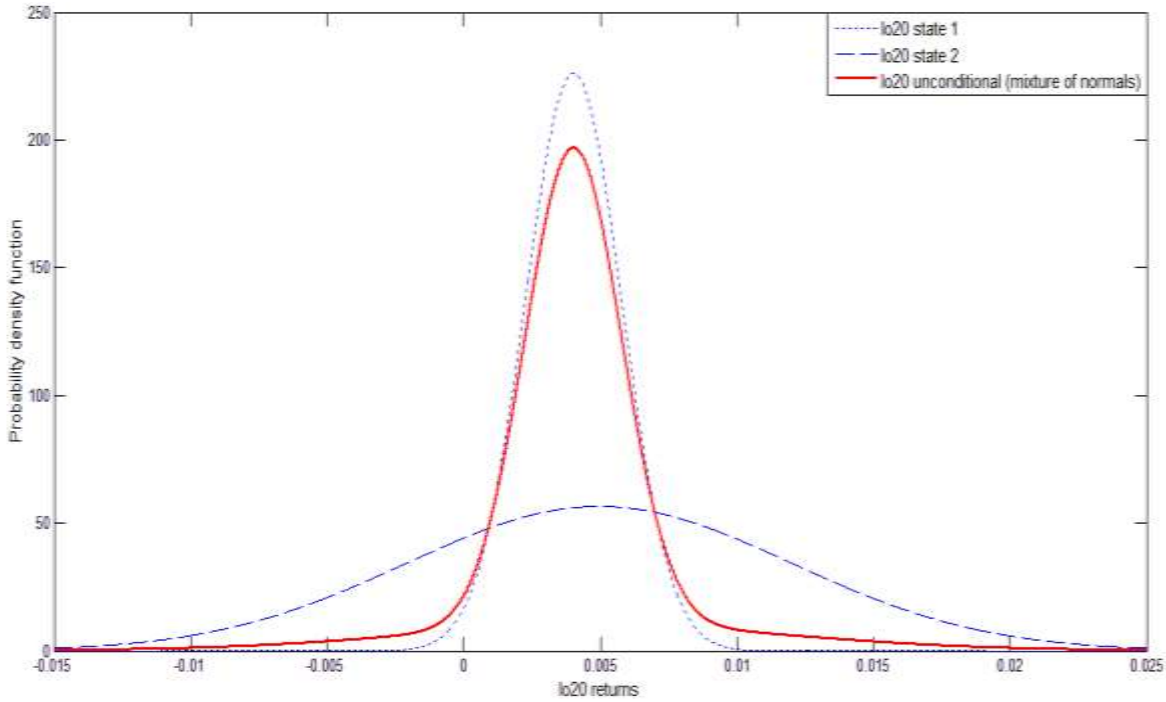


Figure 65 large stocks state conditional and unconditional estimated density function

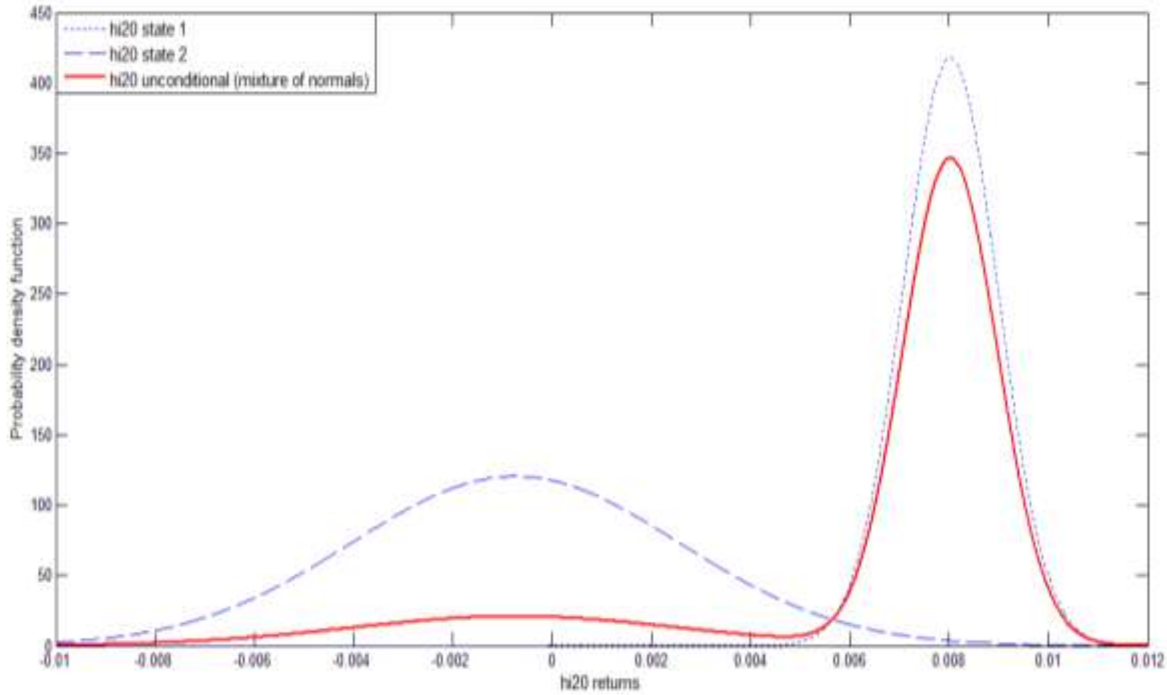
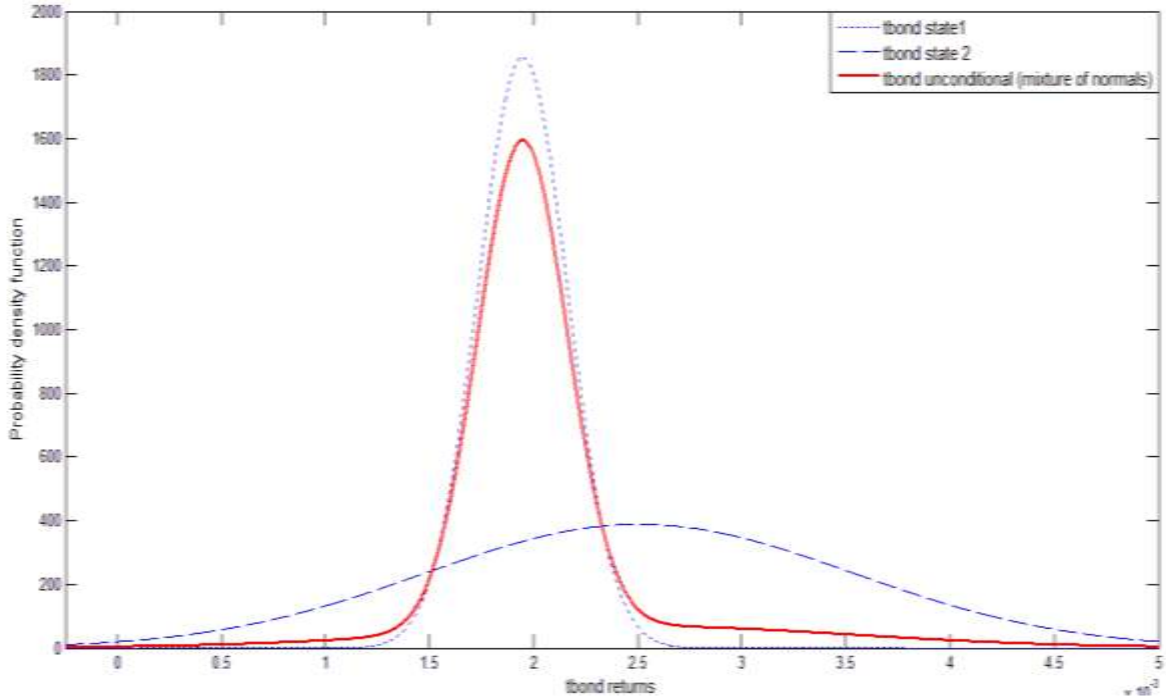


Figure 66 bonds state conditional and unconditional estimated density function



As pointed out by Ang and Timmermann (2012) a natural question might be which portfolios should be optimally held in each regime, and whether there is an optimal portfolio to hedge against the risk of regime changes. The first paper to examine asset allocation with regime changes was Ang and Bekaert (2002a), the paper

examines portfolio choice for a small number of countries. They exploited the ability of the regime switching model to capture higher correlations during market downturns and examine the question of whether such higher correlations during bear markets negate the benefits of international diversification. They find there are still large benefits of international diversification. The costs of ignoring the regimes is very large when a risk-free asset can be held; investors need to be compensated approximately 2 to 3 cents per dollar of initial wealth to not take into account regime changes. The model I have estimated is characterized by two regimes: the high volatility regime has the lowest Sharpe ratio and its mean-standard deviation frontier is the red one while the low volatility regime has the highest Sharpe ratio and its mean-standard deviation frontier is the green one. The unconditional mean-standard deviation frontier averages across the two mean-standard deviation frontiers and is the blue one and it has been constructed using the asset classes unconditional moments. The state 1 conditional efficient frontier is shown as the green line while the state 2 conditional efficient frontier is shown as the red line. As pointed out by Ang and Timmermann (2012) an investor who ignores regimes sits on the unconditional frontier, thus an investor can do better by holding a higher Sharpe ratio portfolio when the low volatility regime prevails. Conversely, when the bad regime occurs, the investor who ignores regimes holds too high a risky asset weight. She would have been better off shifting into the risk-free asset when the bear regime hits. Clarke and de Silva 1998 stated that the presence of two regimes and two frontiers means that the regime switching investment opportunity set dominates the investment opportunity set offered by one frontier. In other words as affirmed by Ang e Bekaert (2002) portfolio allocations based on regime switching estimates have the potential to out-perform because they set up a defensive portfolio in the bear regime that hedges against high correlations and low returns.

Figure 67 regime-dependent mean-standard deviation efficient frontiers

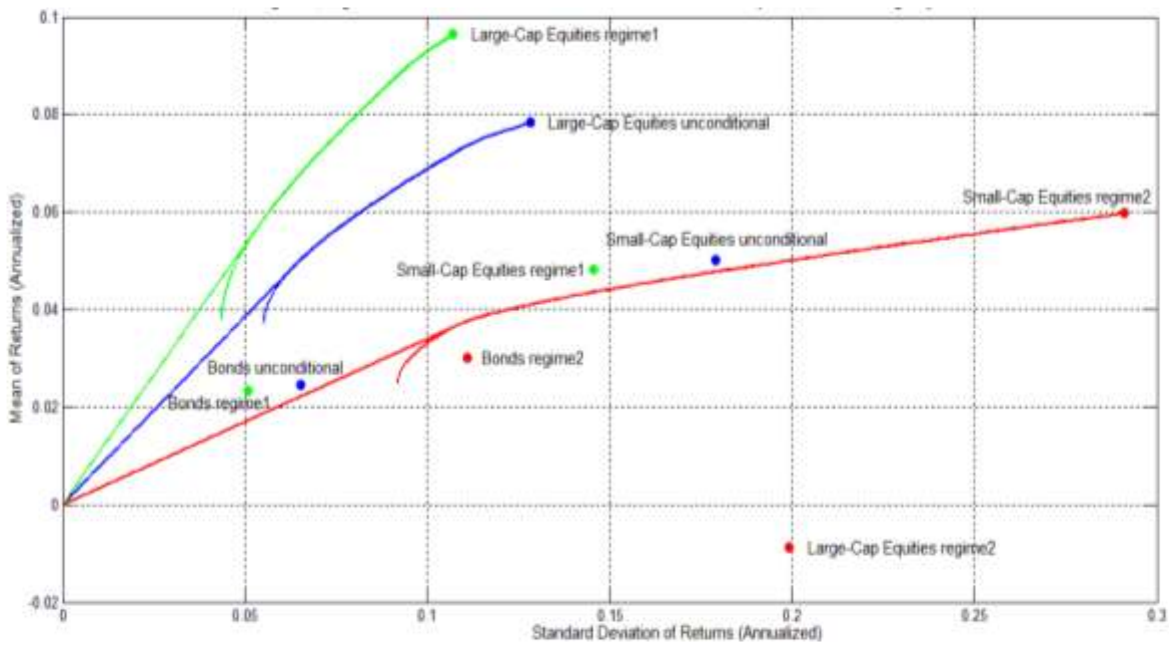


Figure 67 shows the unconditional and the regime-dependent mean-standard deviation efficient frontiers of portfolios that consist of Small-Cap Equities, Large-Cap Equities and bonds. Then means and variances have been annualized. The intercepts of the tangency lines are equal to zero because the vertical axis unit of the Figure 67 is expressed as excess returns over risk-free rate, alike the asset classes returns, consequently the excess return of the risk-free rate is equal to zero. Again the efficient frontiers have been estimated from the state conditional and unconditional asset classes density functions which are characterizes by an expected unconditional mean estimates vector equal to the first three elements of the following vector:

(48)[16] $\mathrm{E}(\boldsymbol{y}_{t+1}|\boldsymbol{\theta}) = \xi_{1\infty}\overline{\boldsymbol{\mu}_1} + (1-\xi_{1\infty})\overline{\boldsymbol{\mu}_2}$

and state conditional mean estimates vectors equal to:

(49)[17] $\overline{\boldsymbol{\mu}_1} = \mathrm{E}(\boldsymbol{y}_t|\mathrm{S_t}=1;\boldsymbol{\theta}) = (\mathrm{I}_4 - \boldsymbol{A}_{1,1})^{-1}\boldsymbol{\mu}_1$ and

(50)[18] $\overline{\boldsymbol{\mu}_2} = \mathrm{E}(\boldsymbol{y}_t|\mathrm{S_t}=2;\boldsymbol{\theta}) = (\mathrm{I}_4 - \boldsymbol{A}_{1,2})^{-1}\boldsymbol{\mu}_2$

[16, 17, 18] Massimo Guidolin "Modelling, Estimating and Forecasting Financial Data under Regime (Markov) Switching" download at "http://didattica.unibocconi.it/mypage/dwload.php?nomefile=Lecture_7_-_Markov_Switching_Models20130520235704.pdf"

while the unconditional covariance matrix adjusted for the regime structure is equal to the *(3 x 3)* left-upper part of the following matrix:

(51)[19] $\mathrm{Var}(\mathbf{y_{t+1}}|\boldsymbol{\theta}) = \xi_{1\infty}\boldsymbol{\Sigma_1} + (1-\xi_{1\infty})\boldsymbol{\Sigma_2} + \xi_{1\infty}(1-\xi_{1\infty})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})(\overline{\boldsymbol{\mu_1}} - \overline{\boldsymbol{\mu_2}})^{\mathrm{T}}$

and state conditional covariance matrices equal to $\boldsymbol{\Sigma_1}$ and $\boldsymbol{\Sigma_2}$.

4.2 In-sample asset allocation exercise

In this paragraph I am going to illustrate the in-sample portfolio construction processes that have been adopted and both the realized and the expected asset allocation and portfolio statistics, using firstly the MSVAR(2,1) and secondly the VAR(1) to model the first two moments of the asset classes returns distributions.

4.2.1 In-sample asset allocation exercise based on a MSVAR(2,1) model

The regime switching, state conditional, efficient frontiers illustrated in Figure 67 are the actual efficient frontiers in which the investor sits only if she knows which regime applies at each time. Given that the model I have estimated is characterized as having a state variable driven by a hidden Markov process, the regimes cannot be identified in real time. Then, the underlying regime is treated as a latent variable that is unobserved and an agent can learn about regimes employing a filtering algorithm. Since the return distribution is very different in the bull and bear state, the state probability perceived by investors is a key determinant of their asset holdings. The filtering algorithm uses a Bayesian rule to update beliefs according to how likely new observations are drawn from different regimes, which are weighted by prior beliefs concerning the previous regimes. The higher the persistence of the regimes, the greater the weight on past data.

In Figure 67, the risk-return trade-offs of each single regime are known because the first two moments of the asset classes returns distributions have been

[19] Massimo Guidolin "Modelling, Estimating and Forecasting Financial Data under Regime (Markov) Switching" download at "http://didattica.unibocconi.it/mypage/dwload.php?nomefile=Lecture_7_-_Markov_Switching_Models20130520235704.pdf"

estimated by the model. However, as pointed out by Ang and Timmermann (2012), an investor who knows the parameters can infer which regime prevails at each time. Then the updating of the probability of the current regime, given all information up to time t, can be computed using methods similar to a learning problem. The investor infers the regime probability from the current information, in particular she computes the probability that a certain regime prevails at the current time, which can be easily computed as a by-product of the estimation of the regimes switching model. This regime probabilities are called filtered probabilities. A filtered probability is the best inference of the current state, based on real time information. Given these probabilities an investor is able to build the filtered dynamic variance-covariance matrix and the filtered dynamic conditional mean vector.

Given that the investor is not interested in invests in the dividend yield but only in the small stocks, large stocks and bonds, the filtered dynamic conditional mean for each *t* is a *(3 x 1)* vector equal to the first three elements of the following vector:

(52) $$\mathrm{E}(\boldsymbol{y}_{\mathbf{t}}|\boldsymbol{y}_{\mathbf{t-1}};\boldsymbol{\theta};\Im_{\mathrm{t}}) = \widehat{\boldsymbol{\mu}}_{\mathrm{t|t-1}} = \xi_{1\mathrm{t}}\mathrm{E}(\boldsymbol{y}_{\mathbf{t}}|\boldsymbol{y}_{\mathbf{t-1}};\mathrm{S_t}=1;\boldsymbol{\theta};\Im_{\mathrm{t}}) + \xi_{2\mathrm{t}}\mathrm{E}(\boldsymbol{y}_{\mathbf{t}}|\boldsymbol{y}_{\mathbf{t-1}};\mathrm{S_t}=2;\boldsymbol{\theta};\Im_{\mathrm{t}})$$

and

(53) $$\mathrm{E}(\boldsymbol{y}_{\mathbf{t}}|\boldsymbol{y}_{\mathbf{t-1}};\mathrm{S_t}=j;\boldsymbol{\theta};\Im_{\mathrm{t-1}}) = \widehat{\boldsymbol{\mu}}_{\mathrm{t|t-1};\mathbf{S_t}=\boldsymbol{j}} = \boldsymbol{\mu}_{\mathbf{S_t}=\boldsymbol{j}} + \boldsymbol{A}_{\mathbf{1},\mathbf{S_t}=\boldsymbol{j}}\boldsymbol{y}_{\mathbf{t-1}}$$

for each *t* is a *(4 x 1)* vector that represents the lo20, hi20, tbond and div_y expected returns at time *t* conditional on time *t-1* returns, time *t* state ($\mathrm{S_t}=j$) and all the past information up on time *t-1*. Here $\boldsymbol{\mu}_{\mathbf{S_t}=\boldsymbol{j}}$ is the vector of intercept terms of $\boldsymbol{y}_{\mathbf{t}}$ in state $\mathrm{S_t}=j$, $\boldsymbol{A}_{\mathbf{1},\mathbf{S_t}=\boldsymbol{j}}$ is the matrix of autoregressive and regression coefficients associated with lag 1 in state $\mathrm{S_t}=j$ and $\Im_{\mathrm{t}}$ is all the past information up to time *t* the except the returns at time *t*. In computing the vector of the expected returns at time *t=1*, conditional on time *t-1=0*, I assumed that $\boldsymbol{y}_{\mathbf{0}}$ is a vector of zeros[20] and that $\xi_{10}=\xi_{20}=0.5$.

[20] Massimo Guidolin "Modelling, Estimating and Forecasting Financial Data under Regime (Markov) Switching" download at "http://didattica.unibocconi.it/mypage/dwload.php?nomefile=Lecture_7_-_Markov_Switching_Models20130520235704.pdf"

The filtered dynamic variance-covariance matrix for each *t* is equal to the *(3 x 3)* left-upper part of the following matrix::

(54) $Var(\boldsymbol{y}_{\mathbf{t}}|\boldsymbol{y}_{\mathbf{t-1}};\boldsymbol{\theta};\Im_{\mathrm{t}}) = \xi_{1\mathrm{t}}\,\Sigma_1 + \xi_{2\mathrm{t}}\Sigma_2 + \xi_{1\mathrm{t}}\xi_{2\mathrm{t}}(\widehat{\boldsymbol{\mu}}_{\mathrm{t|t-1};\mathbf{S_t=1}} - \widehat{\boldsymbol{\mu}}_{\mathrm{t|t-1};\mathbf{S_t=2}})(\widehat{\boldsymbol{\mu}}_{\mathrm{t|t-1};\mathbf{S_t=1}} - \widehat{\boldsymbol{\mu}}_{\mathrm{t|t-1};\mathbf{S_t=2}})^{T}$

again here is assumed that $\xi_{10} = \xi_{20} = 0.5$.

Given that $\xi_{1\infty} \neq \xi_{1\mathrm{t}}$ and $\xi_{2\infty} \neq \xi_{2\mathrm{t}}$ then $\mathrm{E}(\boldsymbol{y}_{\mathbf{t}}|\boldsymbol{y}_{\mathbf{t-1}};\boldsymbol{\theta};\Im_{\mathrm{t}}) \neq \mathrm{E}(\boldsymbol{y}_{\boldsymbol{t+1}}|\boldsymbol{\theta})$. I then used the filtered dynamic variance-covariance matrix and the filtered dynamic conditional mean vector to build five in-sample recursive optimal portfolios.

Figure 68 lo20 and its filtered dynamic conditional mean

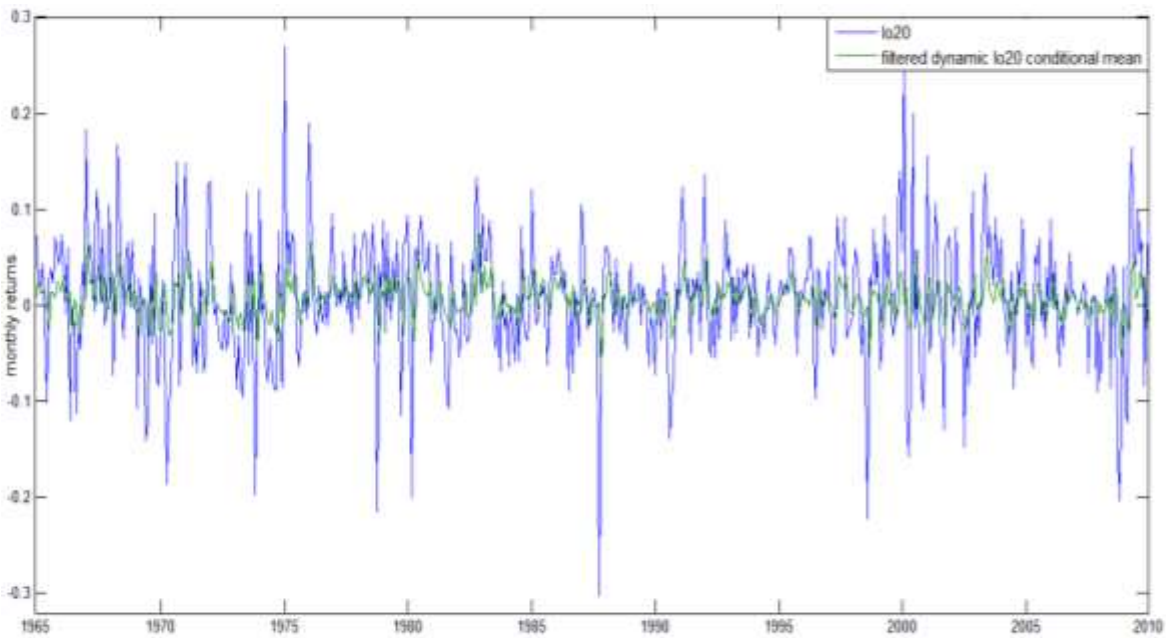


Figure 69 hi20 and its filtered dynamic conditional mean

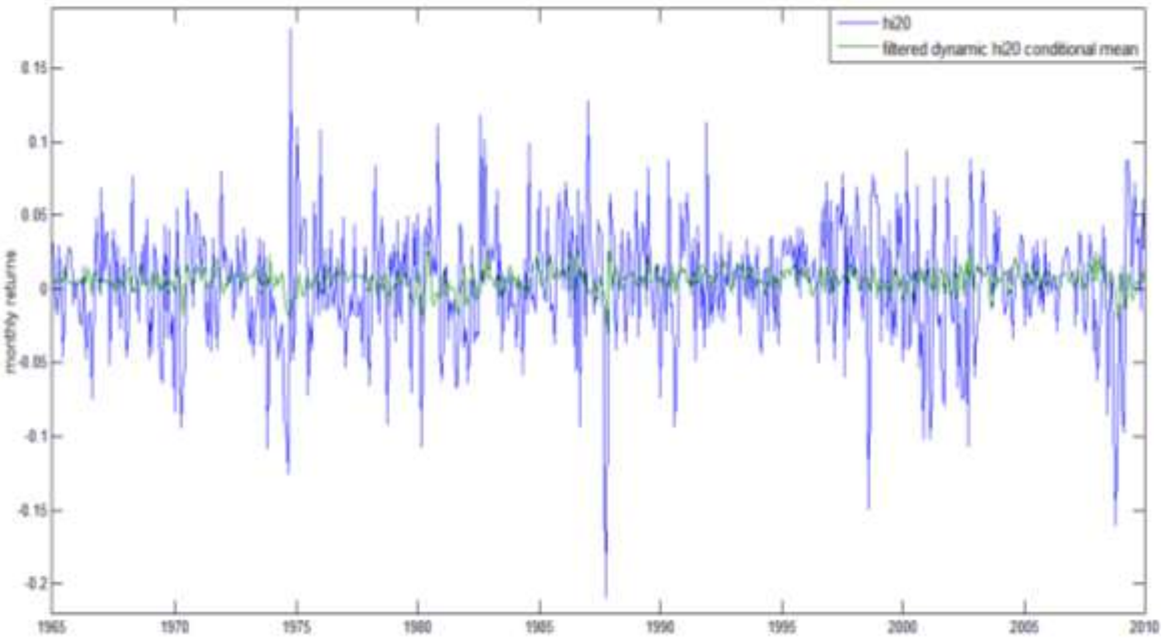
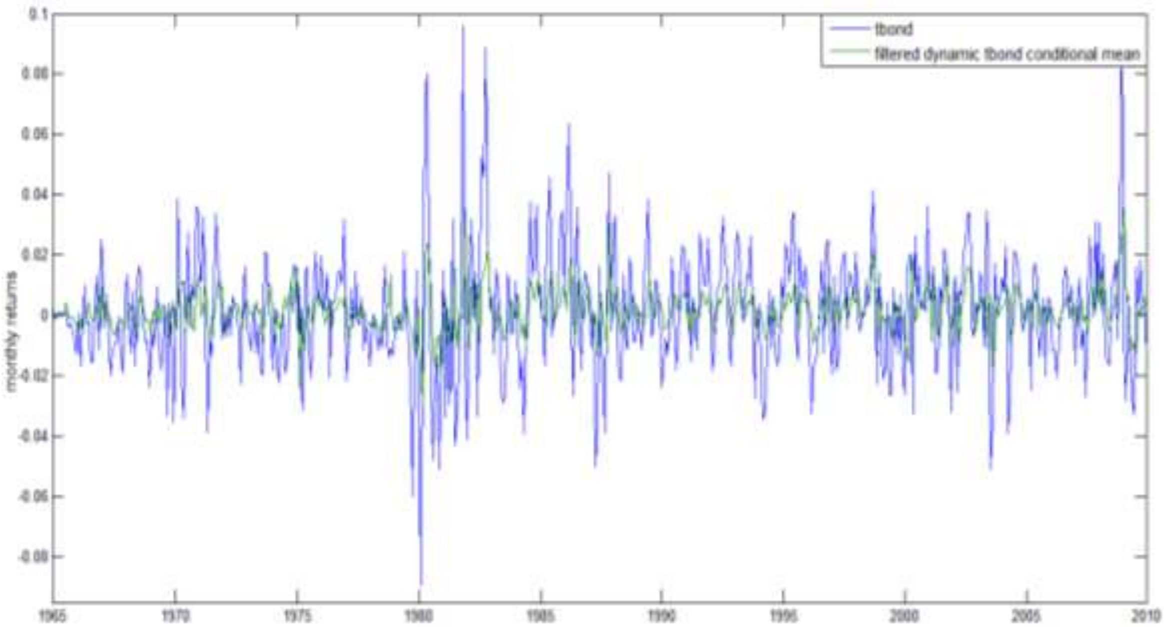


Figure 70 tbond and its filtered dynamic conditional mean



I estimated two in-sample dynamic recursive efficient portfolios that maximize the Sharpe among portfolios on the efficient frontier; in the first in-sample dynamic recursive portfolio the budget constraint is opened up to permit between 0% and 100% in the riskless asset while the second one requires fully-invested portfolios whose weights must sum to 1; in addition short selling, thus negative asset class weights are not allowed. Specifically, a portfolio that maximizes the Sharpe ratio is

also the tangency portfolio on the efficient frontier from the mutual fund theorem; such portfolios are called tangency portfolios since the tangent line from the riskless asset rate to the efficient frontier touches the efficient frontier at portfolios that maximize the Sharpe ratio. The other three in-sample dynamic recursive portfolios have been chosen as those who maximize the investor utility function with three different risk aversion coefficient subject to non negative weights and opened upper budget constraint. One of the factors to consider when selecting the optimal portfolio for a particular investor is degree of risk aversion. This level of aversion to risk can be characterized by defining the investor's indifference curve. This curve consists of the family of risk/return pairs defining the trade-off between the expected return and the risk. It establishes the increment in return that a particular investor requires to make an increment in risk worthwhile. Typical risk aversion coefficients range from 2.0 through 4.0, with the higher number representing lesser tolerance to risk. I choose values of risk aversion coefficients equal t o 1, 3 and 5. I then computed the optimal risky portfolios by generating the efficient frontier from the asset data and then finding the optimal risky portfolio and compute the optimal allocation of funds between the risky portfolio and the riskless asset based on the risk-free rate, the borrowing rate, and the investor's degree of risk aversion. The actual proportion assigned to each of these two investments (the risky portfolio and the riskless asset) is determined by the degree of risk aversion characterizing the investor. If the sum of the computed asset classes weights exceeds 100%, implying that the risk tolerance specified allows borrowing money to invest in the risky portfolio, and that no money is invested in the riskless asset, as a result borrowed capital is added to the original capital available for investment. Tobin's mutual fund theorem (Tobin 1958) says that the portfolio allocation problem can be viewed as a decision to allocate between a riskless asset and a risky portfolio. In the mean-variance framework an efficient portfolio on the efficient frontier serves as the risky portfolio such that any allocation between the riskless asset and this portfolio dominates all other portfolios on the efficient frontier. This portfolio is called a tangency portfolio because it is located at the point on the efficient frontier where a tangent line that originates at the riskless asset touches the efficient frontier. The optimal choice for an investor is the point

of tangency of the highest indifference curve to the Capital Allocation Line CAL, it follows that the slope of the indifference curve is equal to the slope of the CAL.

$$Utility = ExpectedReturn - Variance \cdot 0.005 \cdot RiskAversionCoefficient$$

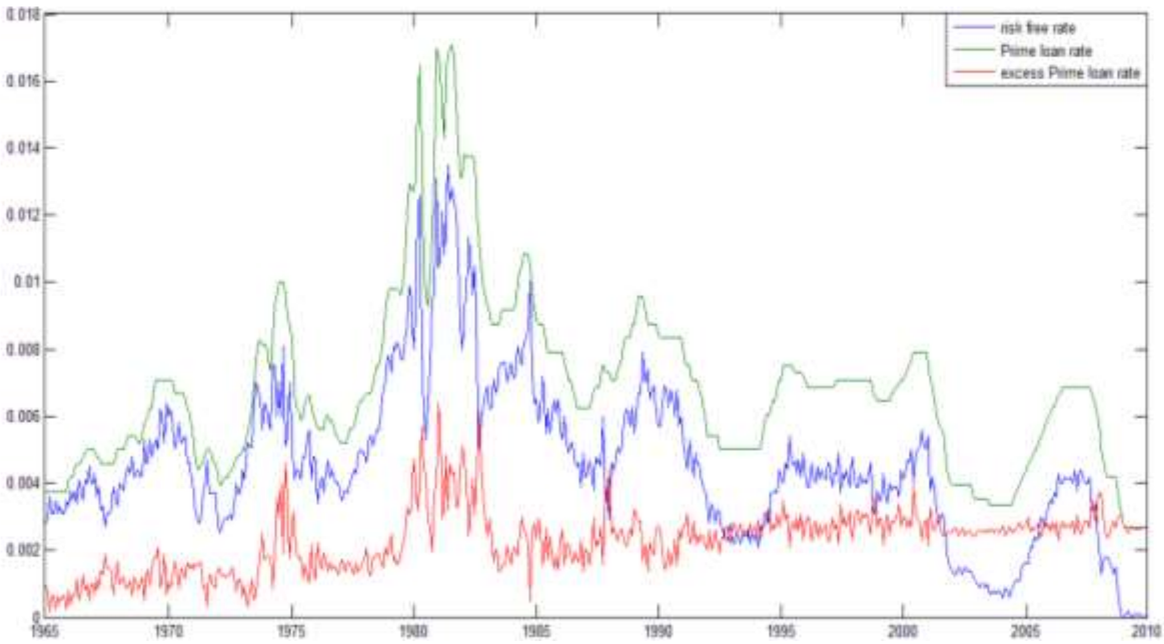
The risk aversion coefficient is a number proportionate to the amount of risk aversion of the investor and is usually set to integer values less than 6, and 0.005 is a normalizing factor to reduce the size of the variance, while variance is the square of the standard deviation, a measure of the volatility of the investment and therefore its risk. This equation is normalized so that the result is a yield percentage that can be compared to investment returns, which allows the utility score to be directly compared to other investment returns. Here is assumed that the riskless asset has variance of 0 and is completely uncorrelated with all other assets. The set of all portfolios with the same utility score plots as a risk-indifference curve. An investor will accept any portfolio with a utility score on her risk-indifference curve as being equally acceptable. Where one of the curves intersects the efficient frontier or the CAL at a single point is the portfolio that will yield the best risk-return trade-off for the risk that the investor is willing to accept. Again, here short selling and negative asset class weights are not allowed. I wrote all the Matlab scripts that carry out the entire computational process.

There are a number of ways to implement a portfolio mean-variance optimization. As already stated, I choose to implement an excess return framework, thus each asset classes returns are expressed as excess returns over the risk-free rate; the risk-free rate is known at each period and is equal to the 1-month US T-Bill monthly returns. Hence, the risk free-rate varies over time as I implement the allocation program. The tangency portfolio, the 100% risky portfolio (the one that invests only in risky assets and not in cash), will hence move over time as well. I also use a borrowing rate represented by the monthly Bank Prime Loan Rate[21] made available by the Board of Governors of the Federal Reserve System (US). It

[21] Board of Governors of the Federal Reserve System (US), Bank Prime Loan Rate [MPRIME], retrieved from FRED, Federal Reserve Bank of St. Louis https://research.stlouisfed.org/fred2/series/MPRIME, February 12, 2016.

is the rate posted by a majority of top 25 (by assets in domestic offices) insured U.S.-chartered commercial banks and it is one of the several base rates used by banks to price short-term business loans. Given that I adopted an excess returns framework also the borrowing rate has been reduced by the amount of the risk free rate, in practice when some money are borrowed to be invested in the risky portfolio then the portfolio excess returns over the risk free (lending) rate are reduced by an amount equal to the excess of the borrowing rate over the risk free (lending) rate; in the time slot analyzed by this study the borrowing rate is always higher than the lending rate (risk free rate), as a consequence borrowing money is always costly in terms of portfolio expected returns.

Figure 71 lending rate (risk free rate), borrowing rate (Prime loan rate) and excess borrowing rate (excess Prime loan rate)



It is noticed that if the risk free lending and borrowing rates are equal, the optimum risky portfolio is obtained by drawing a tangent to the portfolio frontier from the level of risk free lending and borrowing rate, however when the two rates are different the CAL is not unique, one starts form the lending rate and the one that starts from the borrowing rate, in this study, is slightly kinked as the borrowing rate is higher than the risk-free lending rate. If the risky fraction exceeds 1 (100%), implying that the risk tolerance specified allows borrowing money to invest in the risky portfolio, and that no money is invested in the risk-free asset. This borrowed

capital is added to the original capital available for investment. As a consequence the efficient frontier of portfolios with borrowing and lending rate consists of a line segment equal to the CAL that starts from the lending rate for portfolios characterized by a risky fraction smaller than 1, a curve portion of the efficient frontier with neither borrowing nor lending for those portfolios characterized by a risky fraction equal to 1 and finally another line segment equal to the CAL that starts from the borrowing rate for portfolios characterized by a risky fraction greater than 1.

In the following section of the subparagraph the portfolio construction process results are exposed.

Figure 72 maximum Sharpe ratio portfolios weights (1 = 100%) with opened lower budget constraint (permit to invest in the riskless asset)

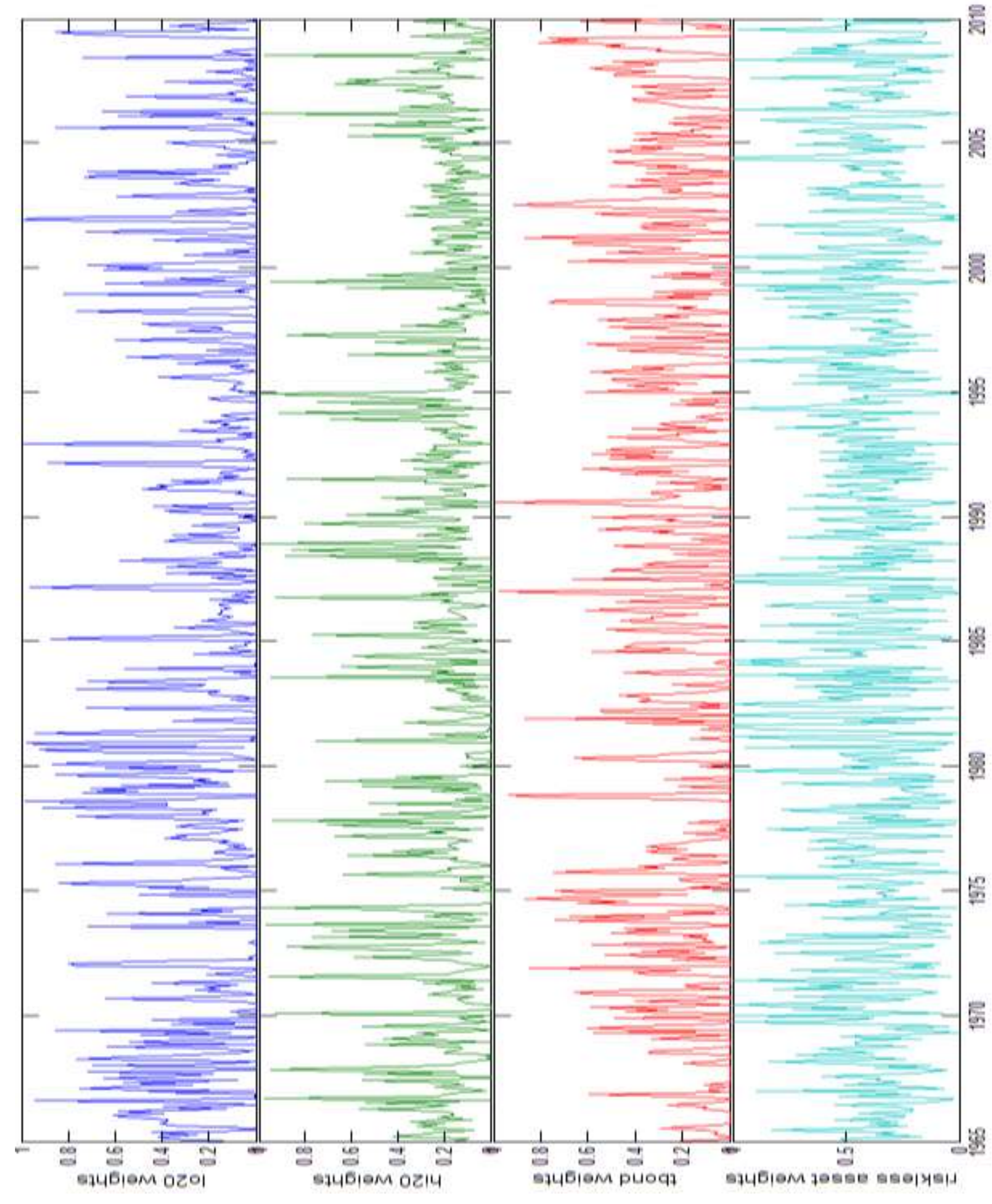


Figure 73 maximum Sharpe ratio portfolios weights (1 = 100%) with budget constraint (not permit to invest in the riskless asset)

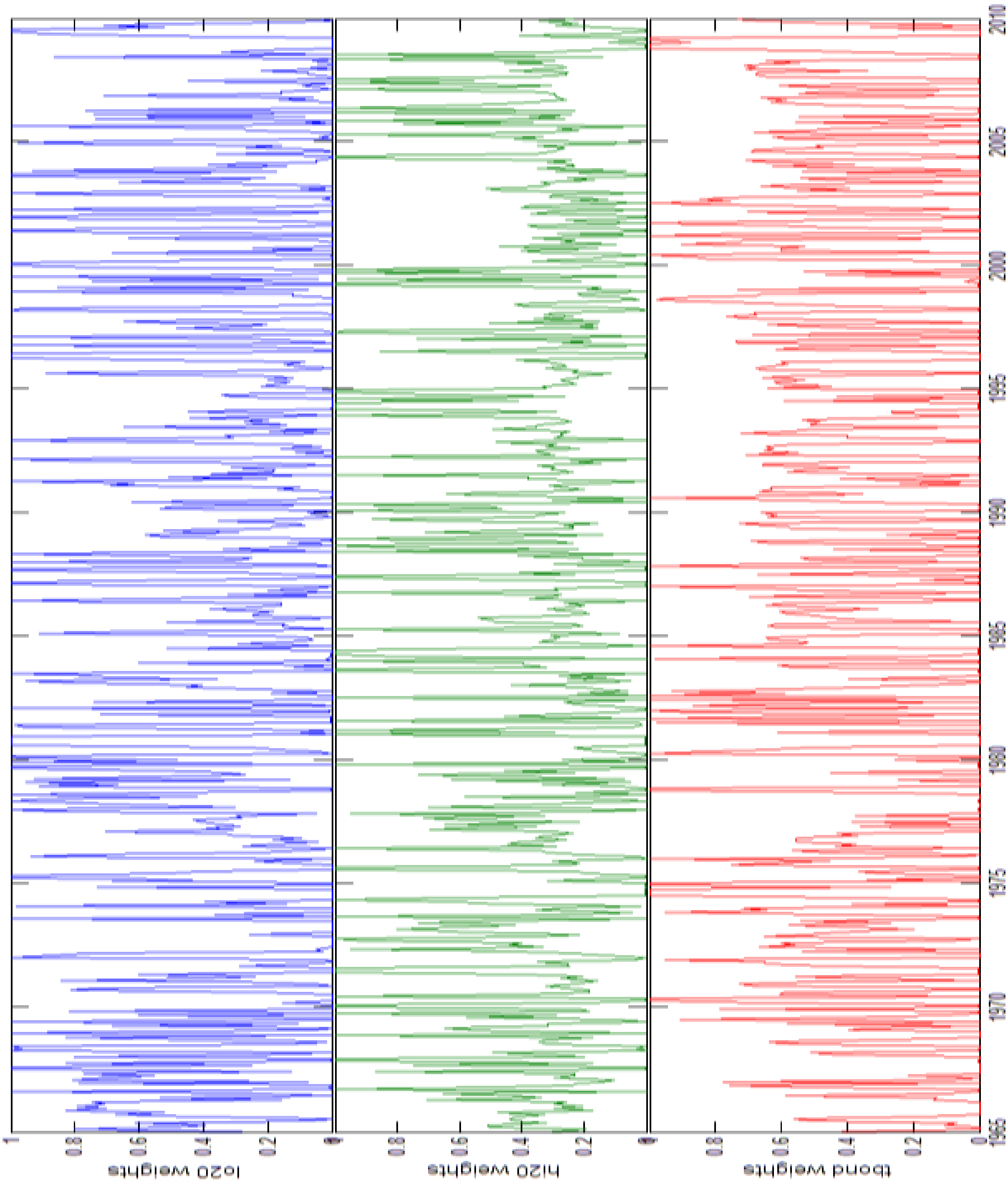


Figure 74 overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 1

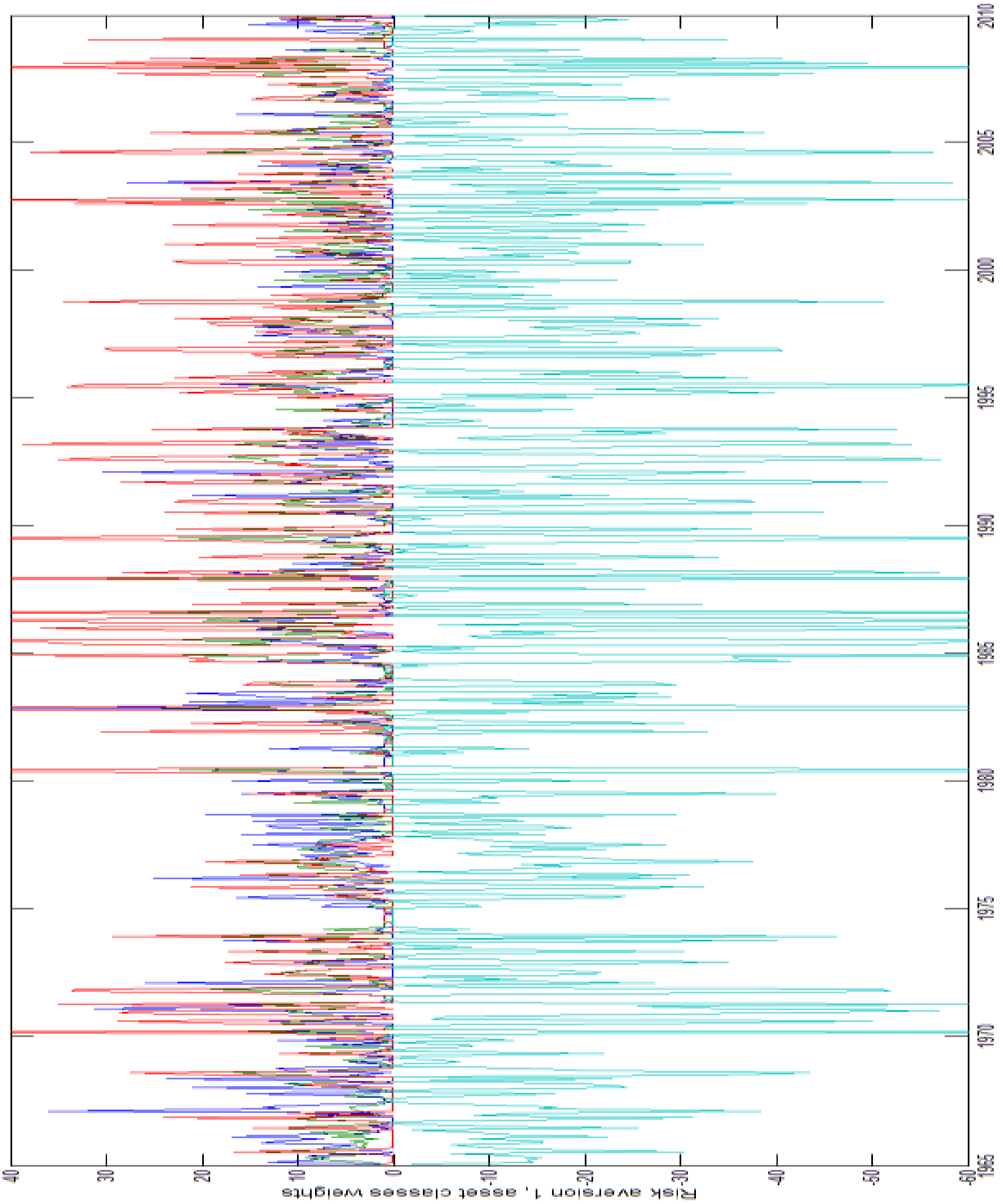


Figure 75 overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 3

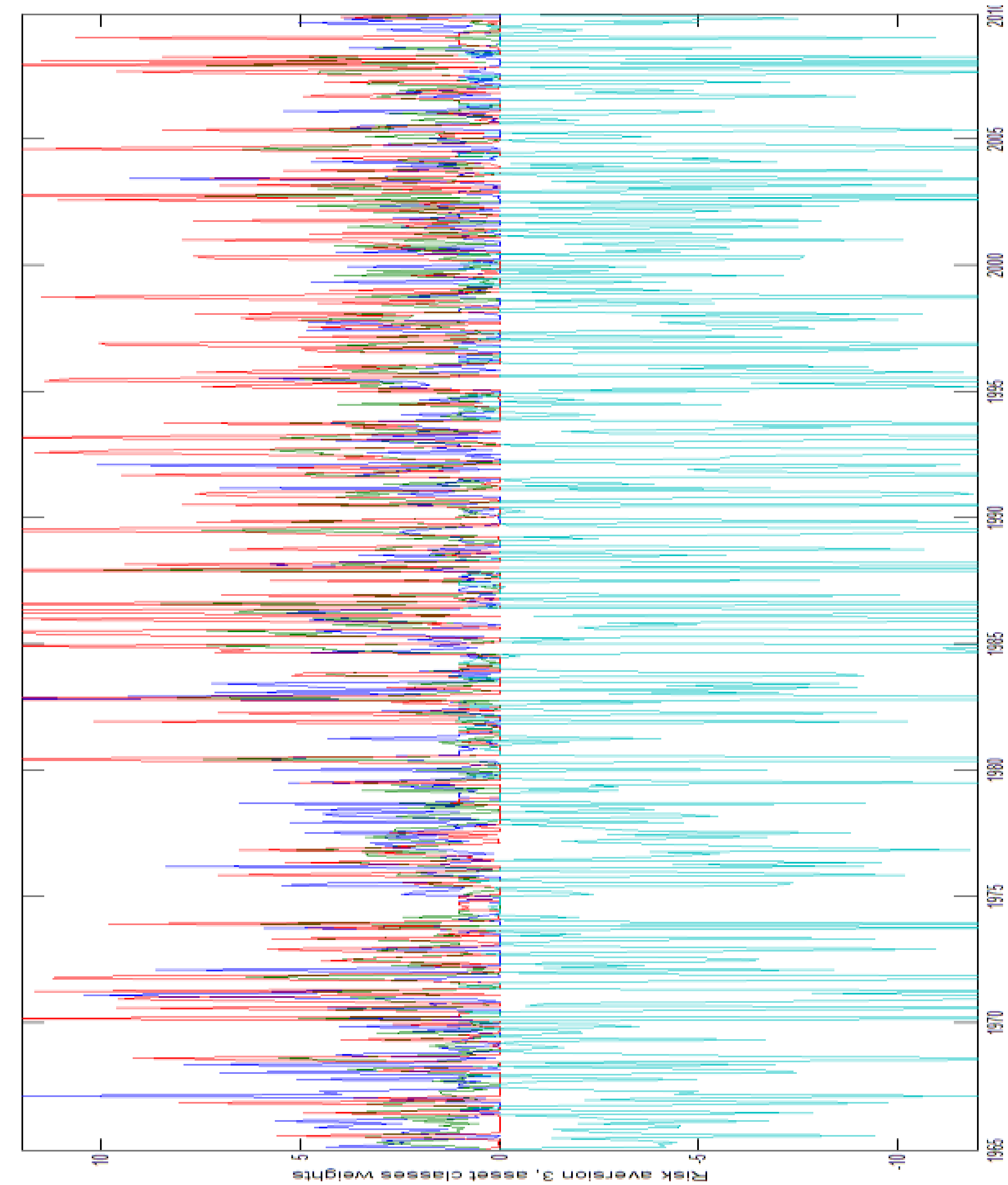
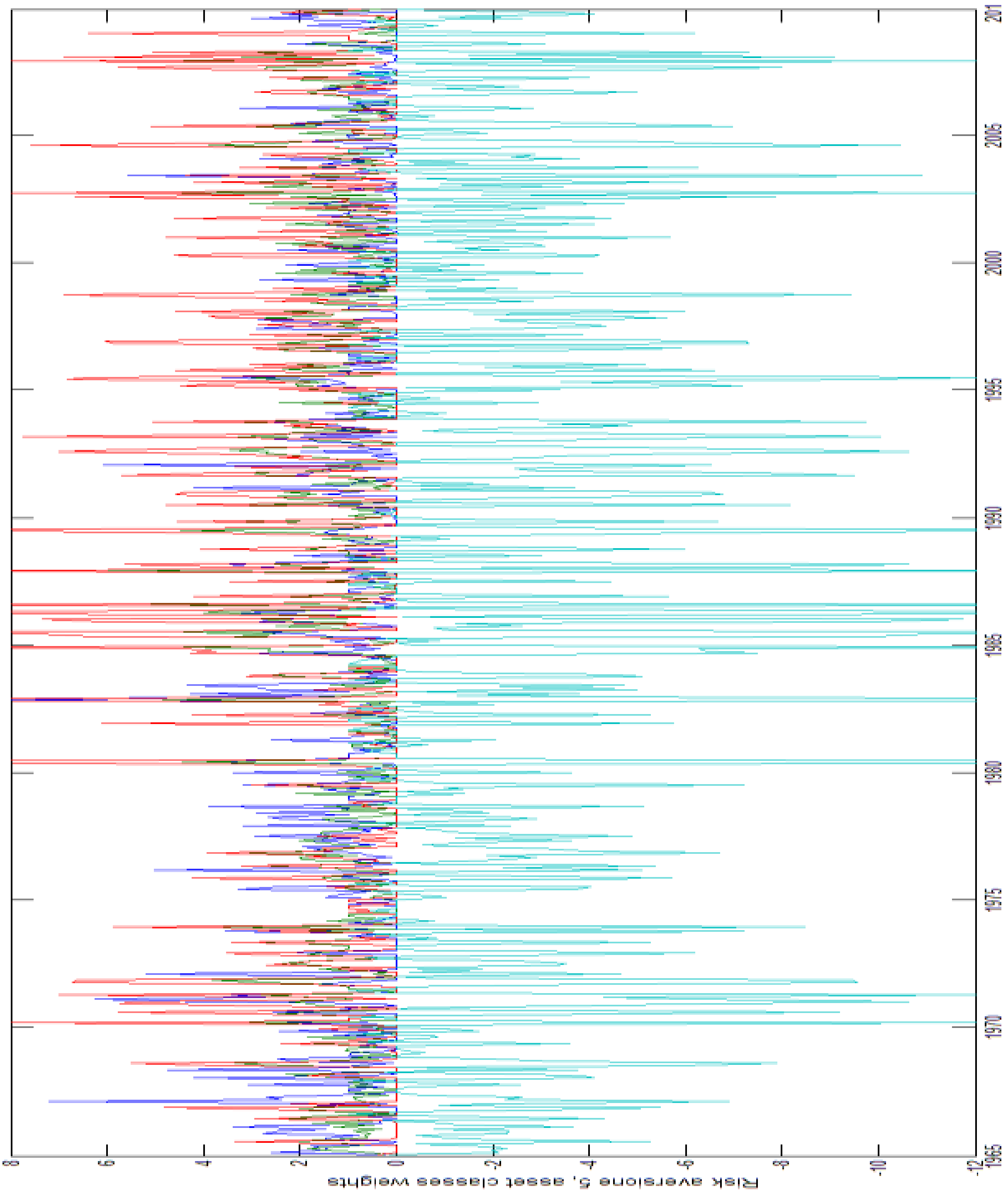


Figure 76 overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 5



Figures 72 to 76 show the overall recursive dynamic weights of the five different built portfolios. As it can be seen the average borrowing factor is very high for the three portfolio built optimizing the expected utility using the capital allocation line with lending and borrowing option, as it might have been largely expected the greater the risk aversion coefficient, the smaller the leverage factor. Conversely, the portfolios built with the first two methods, which is portfolios chosen among those of the efficient frontiers in order to maximize the expected Sharpe ratio (possibly with a share allocated to the riskless asset for the second portfolio construction method), did not resort to any leverage.

Figure 77 risky optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 1

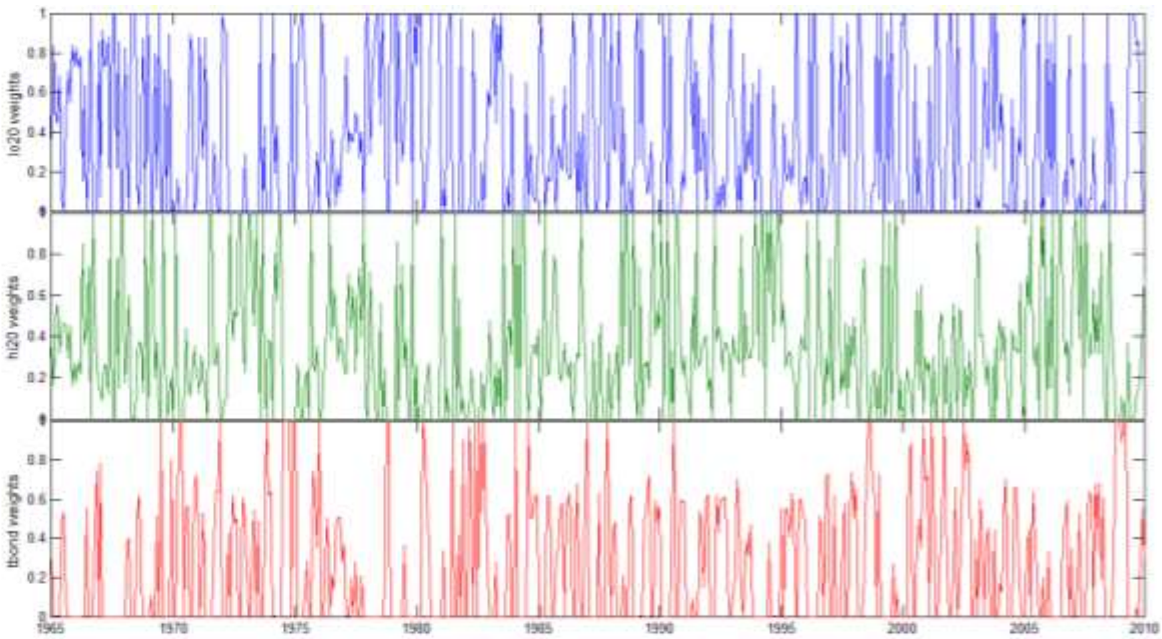


Figure 78 risky optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 3

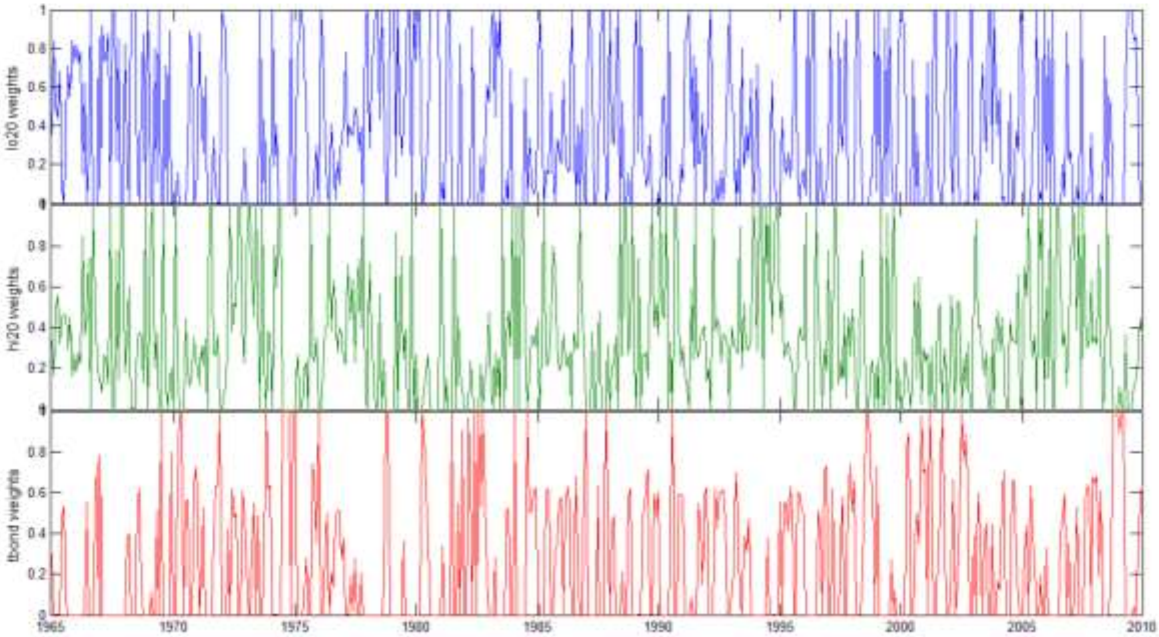
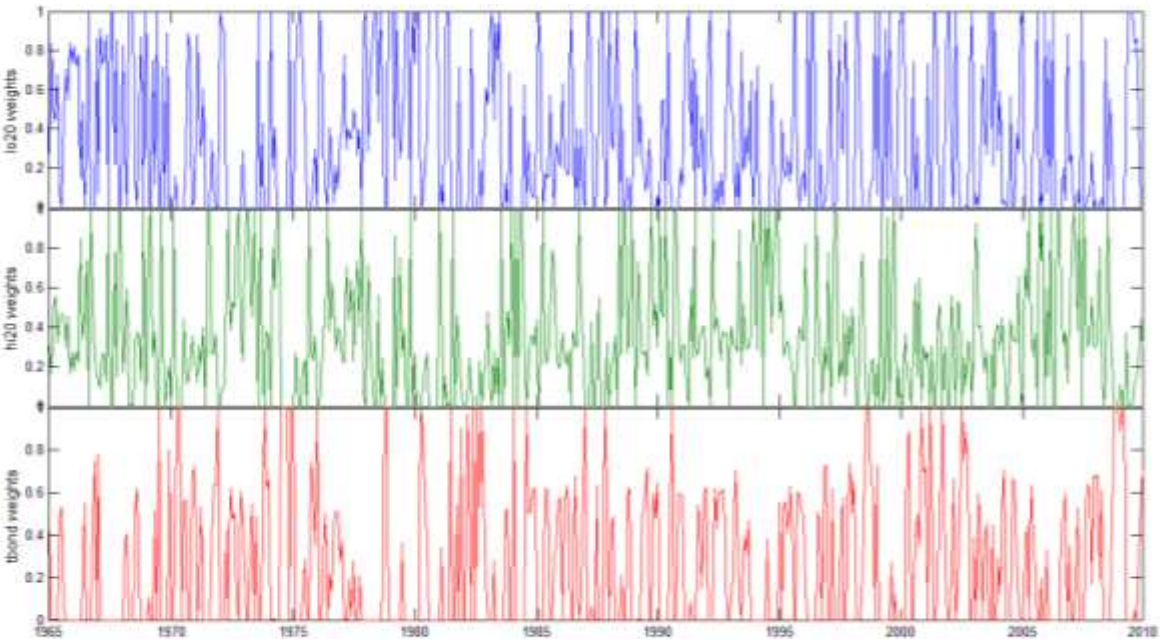


Figure 79 risky optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 5



Figures 77 to 79 show the optimal weights of the tangency portfolios between the capital allocation line (either the lending capital allocation line, or the borrowing capital allocation line) and the efficient frontier.

Figure 80 scatter of the optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with opened lower budget constraint (permit to invest in the riskless asset)

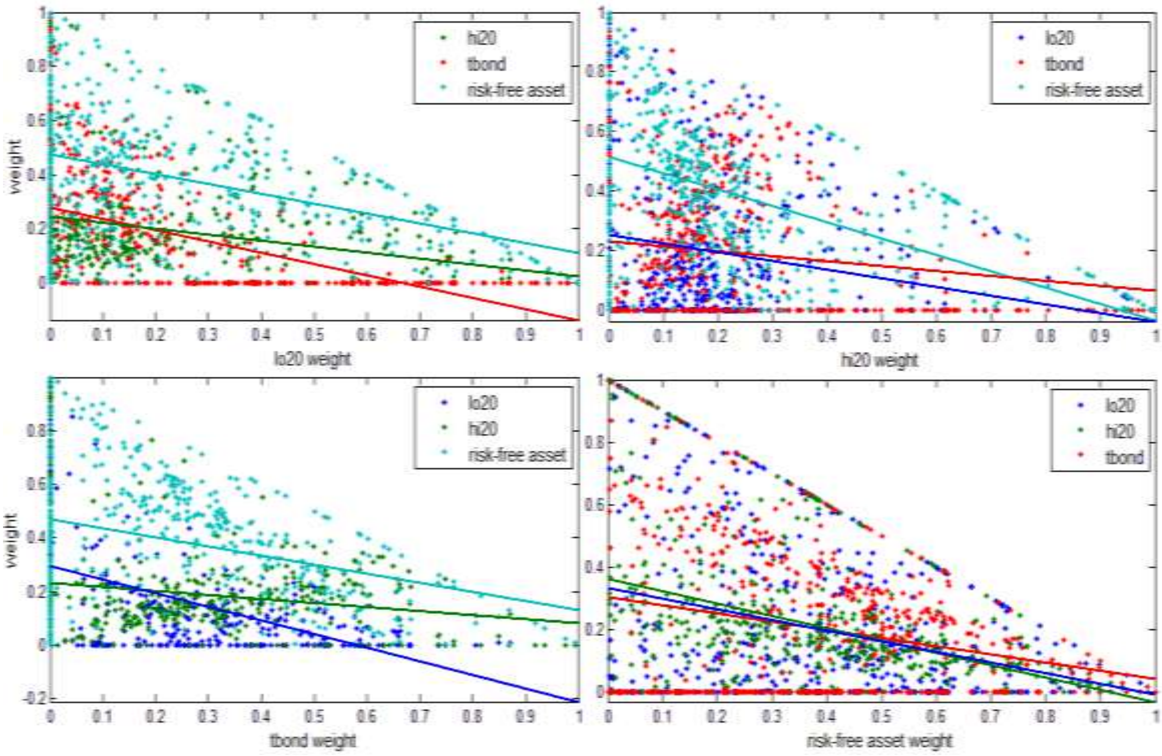


Figure 81 scatter of the optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with budget constraint (not permit to invest in the riskless asset)

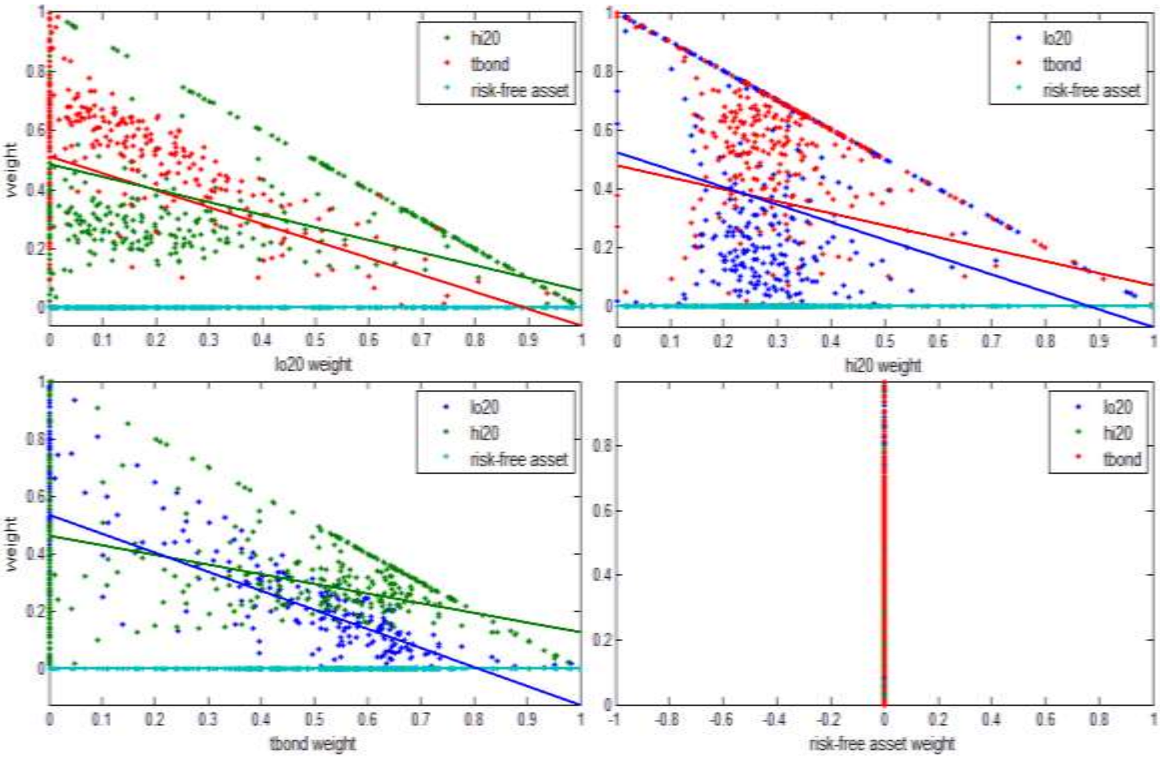


Figure 82 scatter of the optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 1

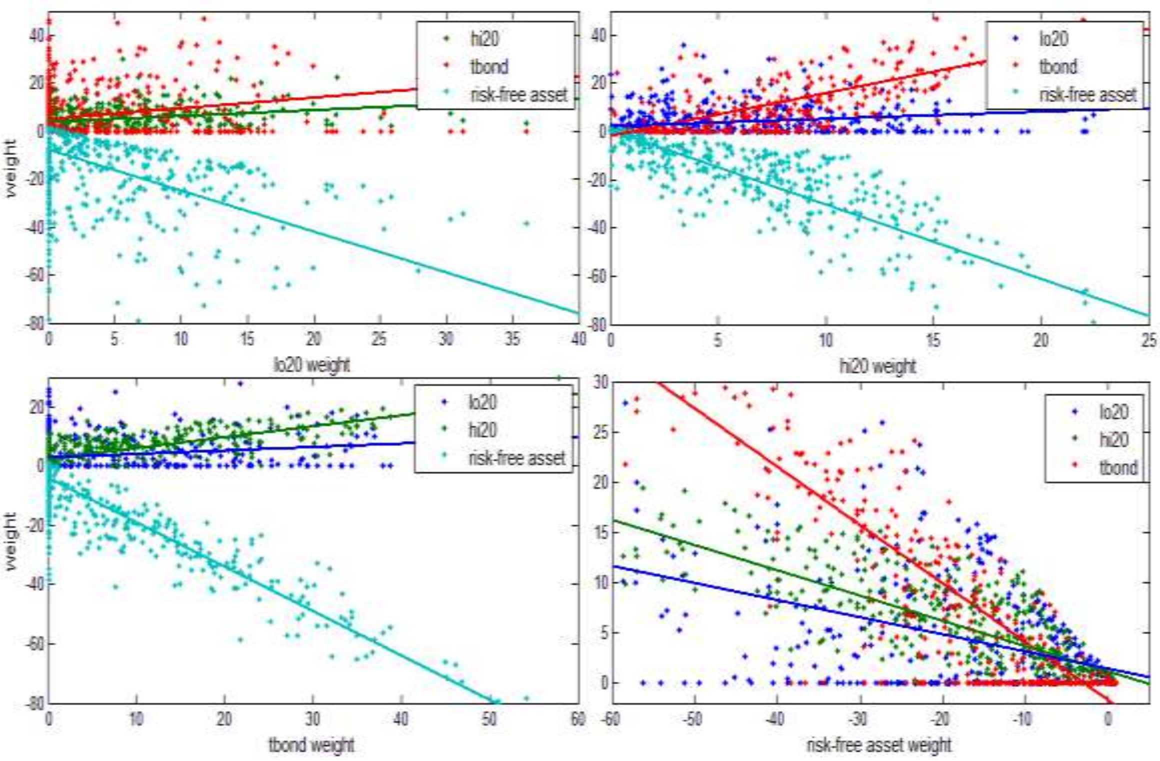


Figure 83 scatter of the optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 3

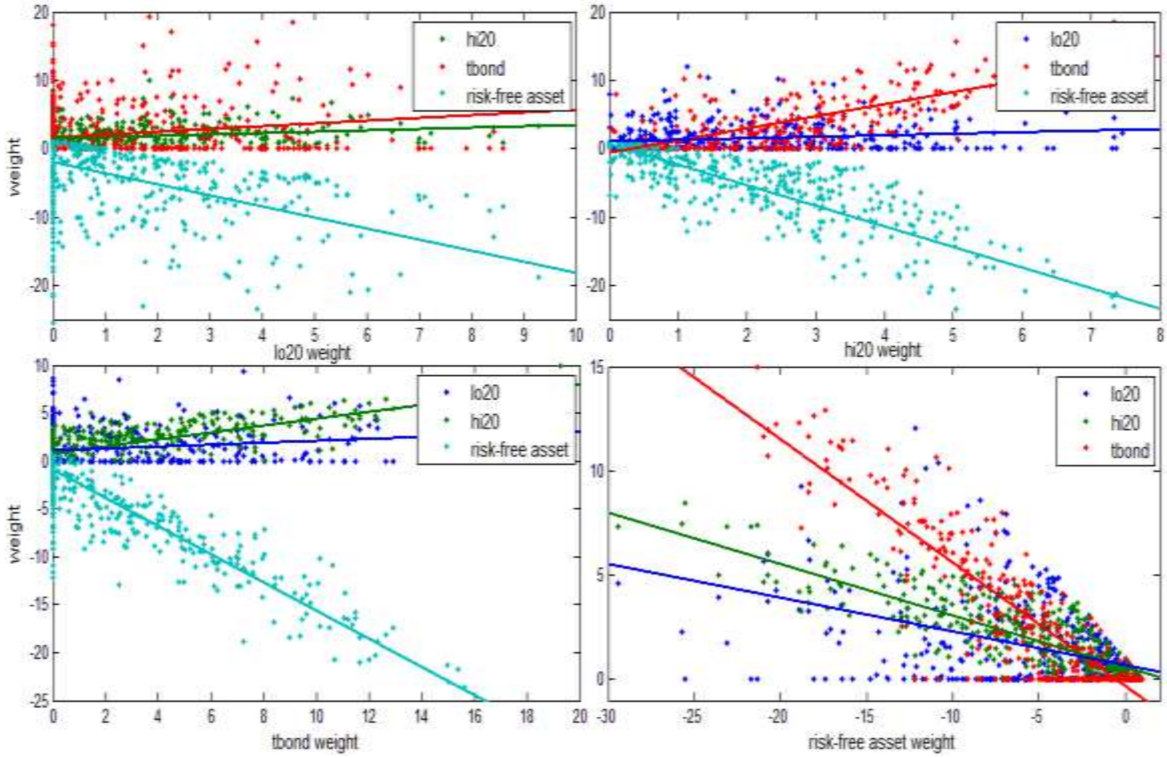
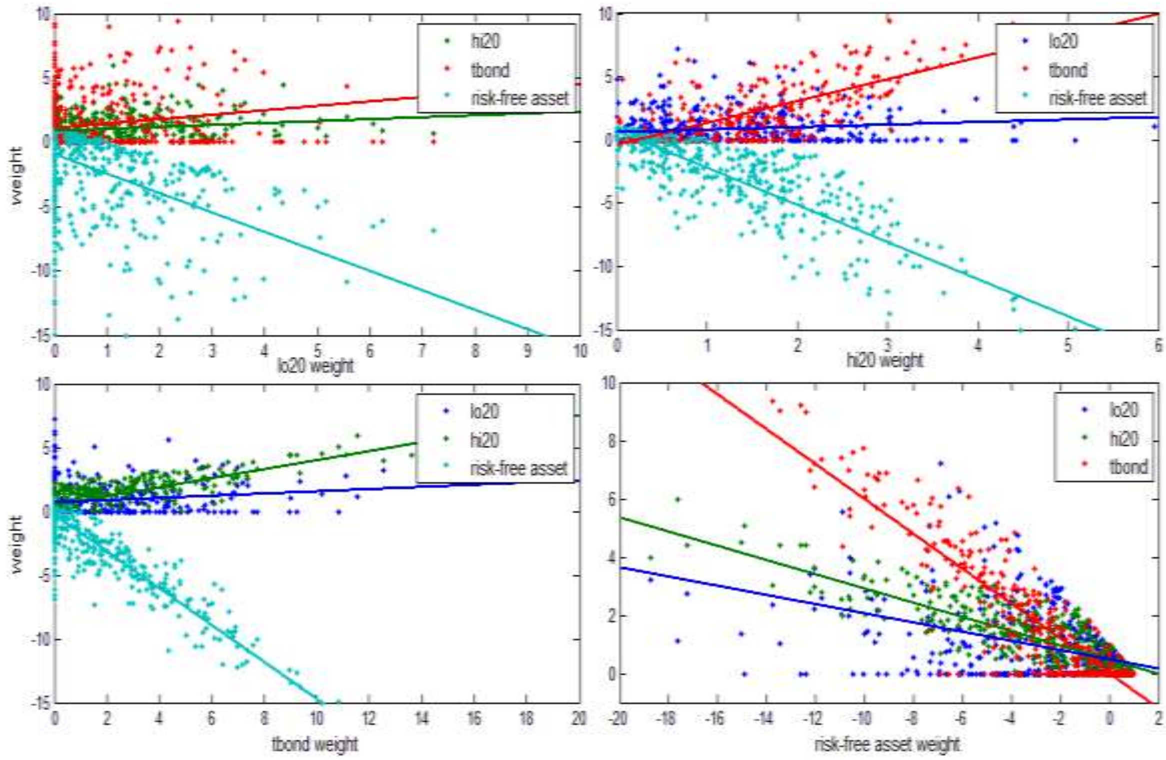


Figure 84 scatter of the optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 5



It is interesting to notice from Figure 80 that is present a sort of substitution effect between lo20 and hi20 optimal weights, when more than half of the portfolio is allocated to the former one, and to a small extent between tbond and hi20 optimal weights when tbond weights are larger than 50%. This relation between lo20 and hi20 optimal weights might be explained by the fact that during bull markets the growth in the lo20 allocation, as long as lo20 weights are low, it is coupled with a reduction of the less profitable asset class, namely the tbond, however when lo20 weights represent a conspicuous share of the portfolios, a further growth of its share in the portfolio is possible only at expense of the most profitable asset class, namely the hi20. Similarly, the substitution effect between tbond and lo20 optimal weights might be originated by volatility reasons since the tbond asset class is the less volatile and the lo20 the most volatile. Volatility can also explain the marked positive relation between hi20 and tbond optimal weights in Figures 82, 83 and 84; weights on those two asset classes jointly grow as the investor flies from the most volatile asset class, namely lo20, to the less volatile asset classes, indeed hi20 and tbond.

Figure 85 scatter of the optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with opened lower budget constraint (permit to invest in the riskless asset) against state 1 smoothed probability

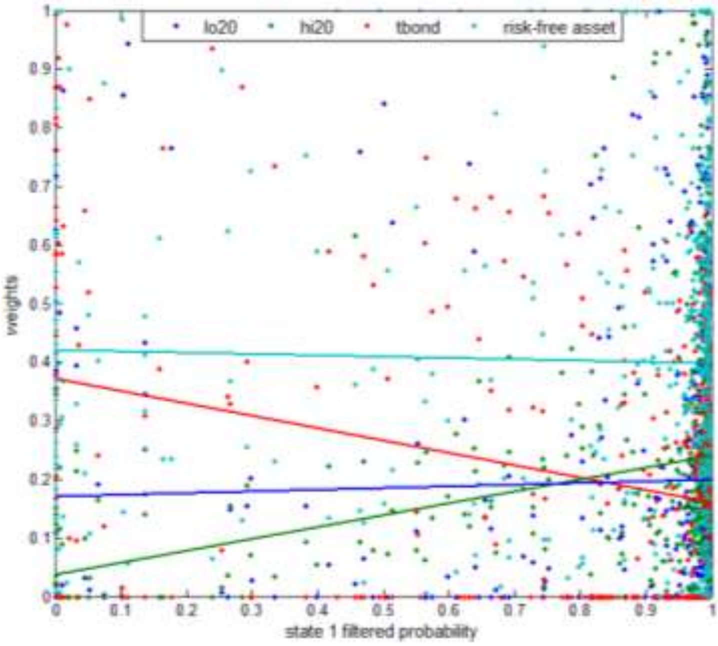


Figure 86 scatter of the optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with budget constraint (not permit to invest in the riskless asset) against state 1 smoothed probability

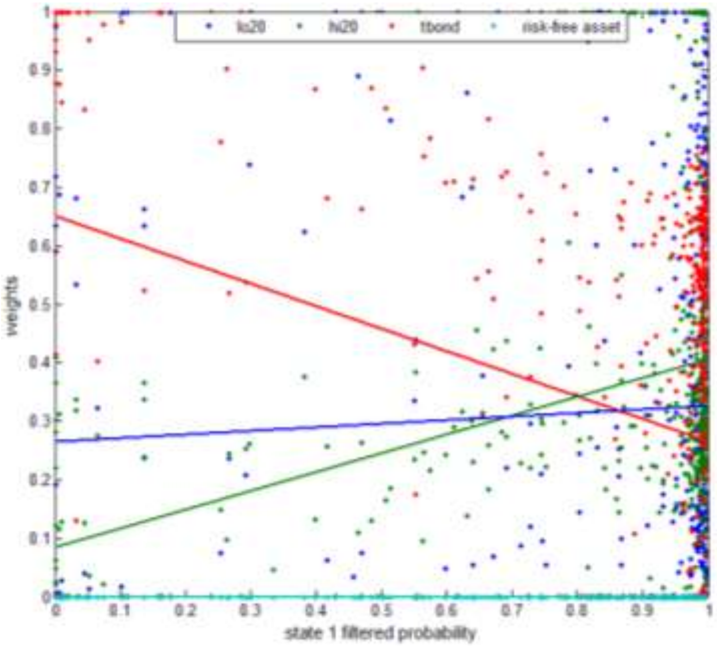


Figure 87 scatter of the optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 1 against state 1 smoothed probability

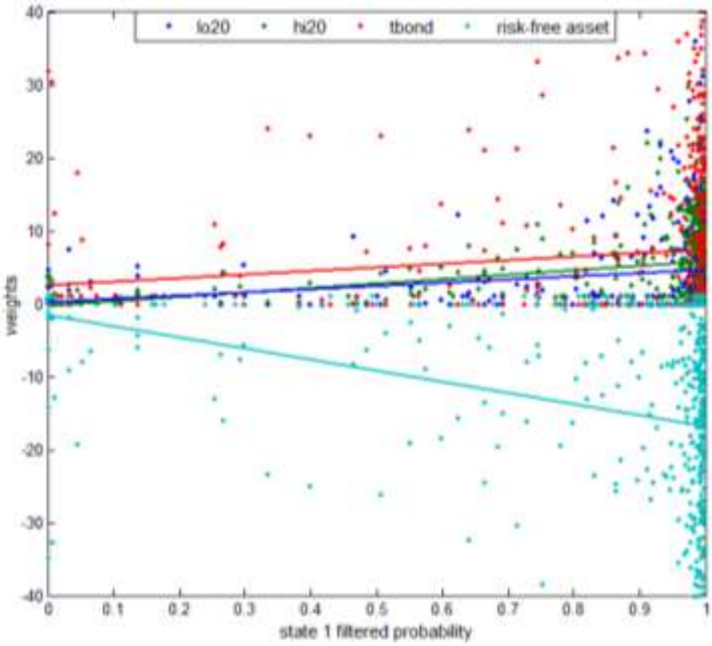


Figure 88 scatter of the optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 3 against state 1 smoothed probability

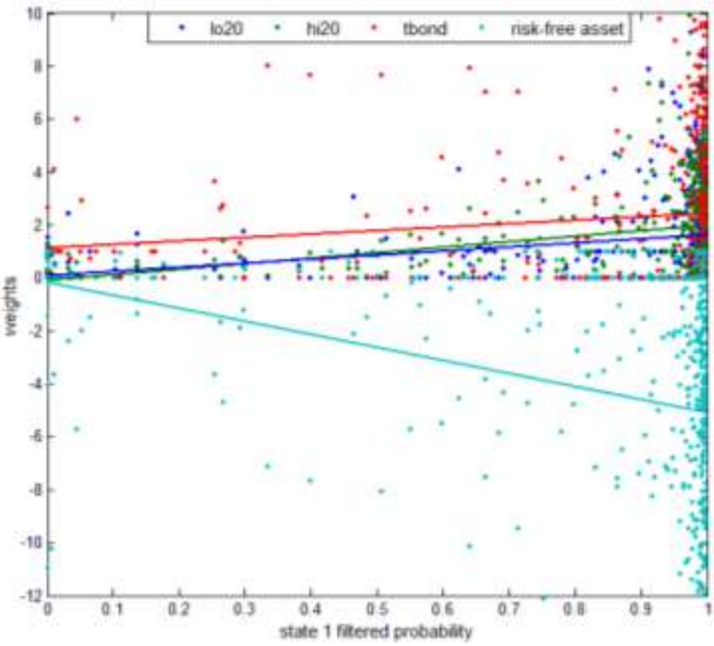
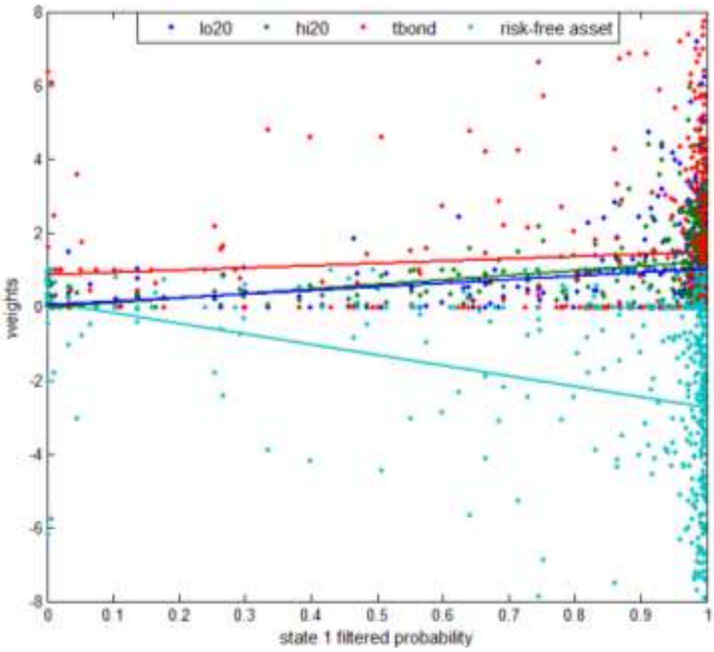


Figure 89 scatter of the optimal weights(1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 5 against state 1 smoothed probability



Figures 85 and 86 shows how lo20 and hi20 optimal weights grow as a function of the perceived probability of being in state 1 since their unconditional means in this state are higher than the tbond ones, which, indeed, decrease as a function of state 1 filtered probability. The same dynamic is less obvious in Figures 87, 88 and 89 because of the leverage factor. Nevertheless, the three figures highlight how more often the investor resorts to borrowing money when he perceives a high probability of being in state 1 rather than in state 2 (low state 1 filtered probability corresponds to high state 2 filtered probability).

Figure 90 scatter of the optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with opened lower budget constraint (permit to invest in the riskless asset) against the dividend yield

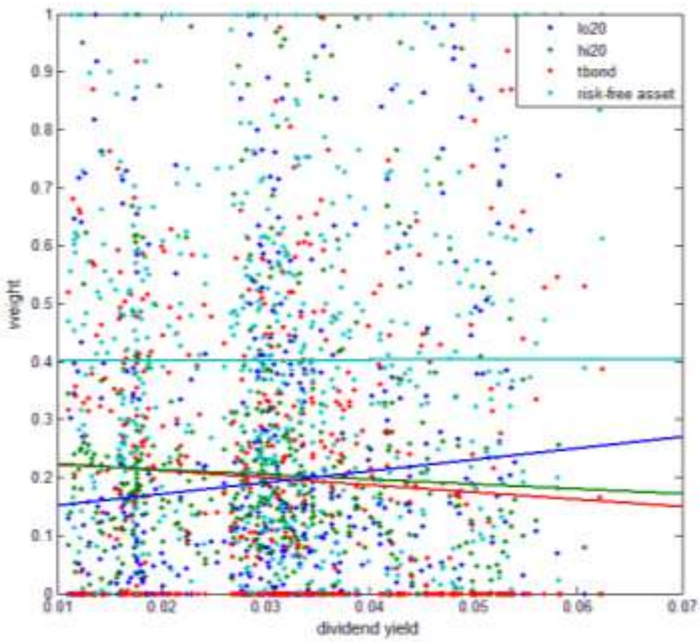


Figure 91 scatter of the optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with budget constraint (not permit to invest in the riskless asset) against the dividend yield

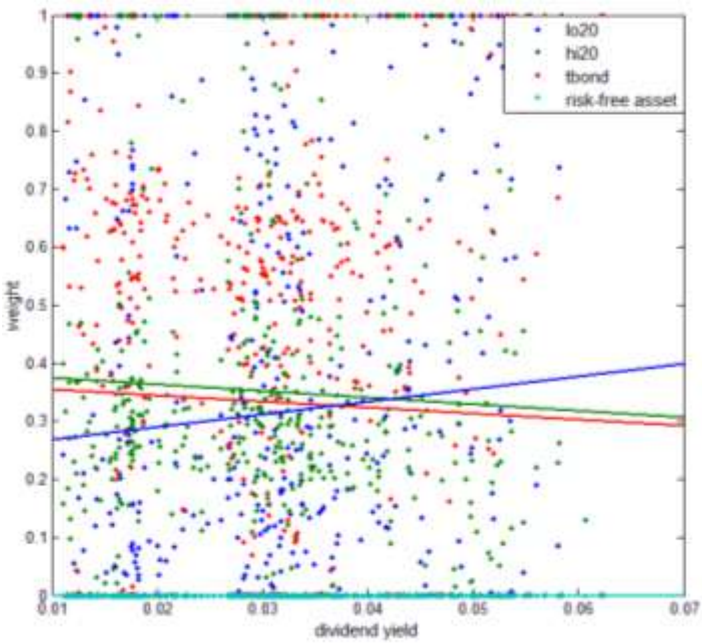


Figure 92 scatter of the optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 1 against the dividend yield

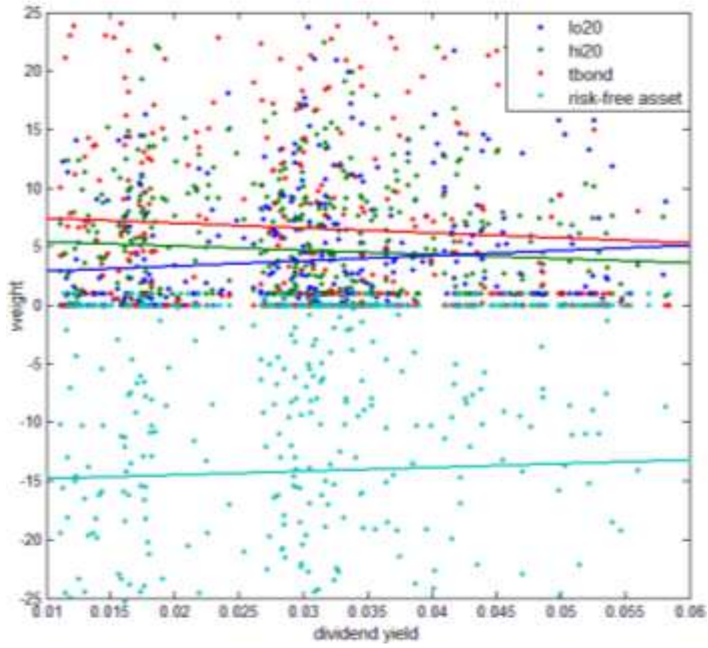


Figure 93 scatter of the optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 3 against the dividend yield

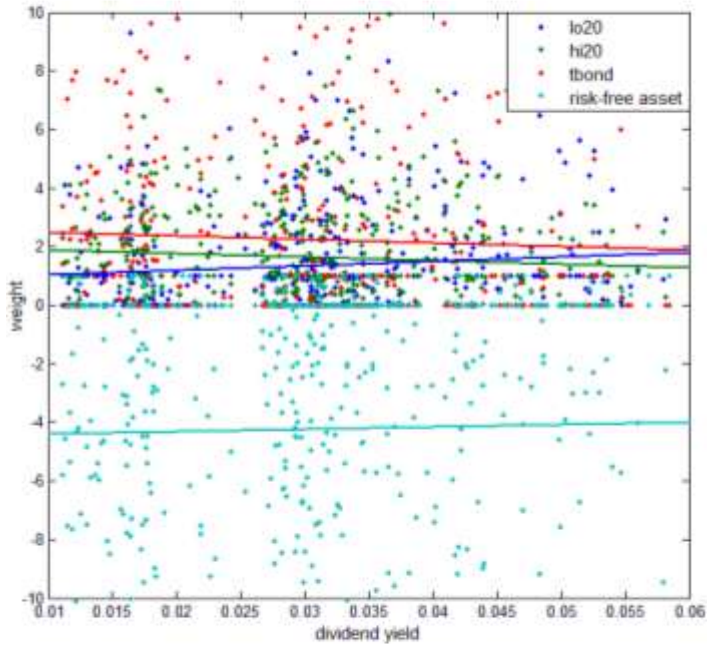
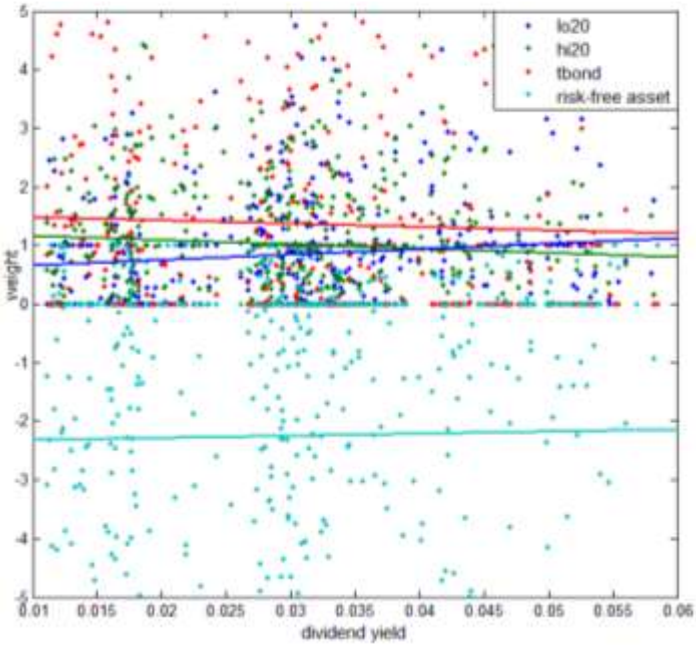


Figure 94 scatter of the optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 5 against the dividend yield



From Figures 90 to 94 it seems that lo20 optimal weights, on the average, increase as a function of the dividend yield while hi20 and tbond optimal weights decrease as a function of that; the relations just discussed however are not that strong.

Figure 95 small stocks optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with opened lower budget constraint (permit to invest in the riskless asset) as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

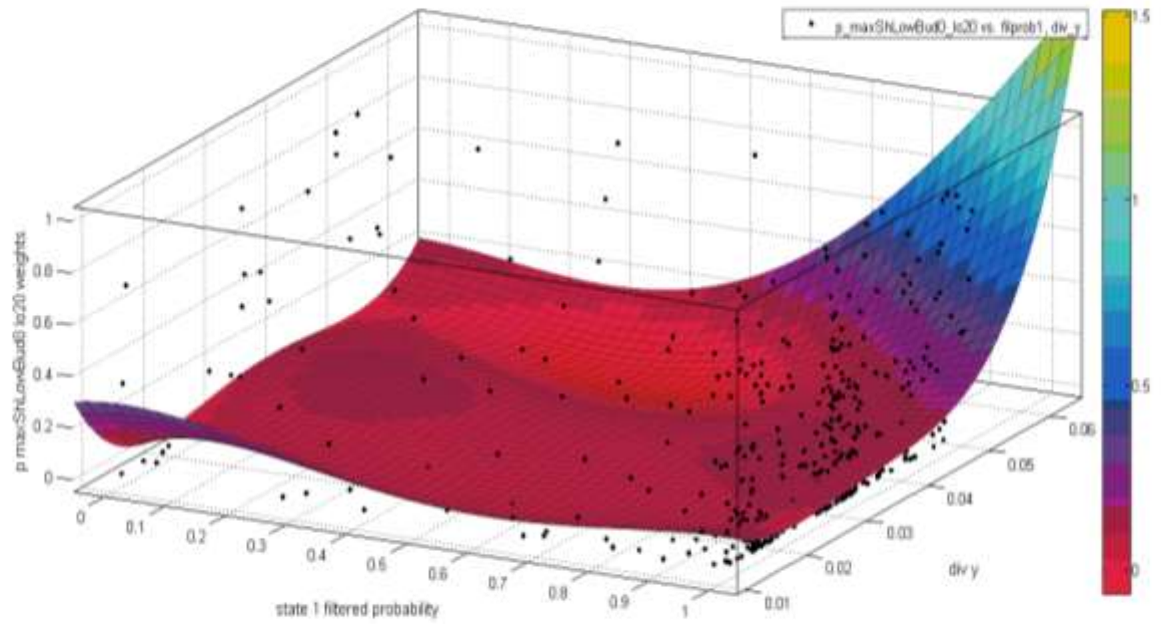


Figure 96 large stocks optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with opened lower budget constraint (permit to invest in the riskless asset) as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

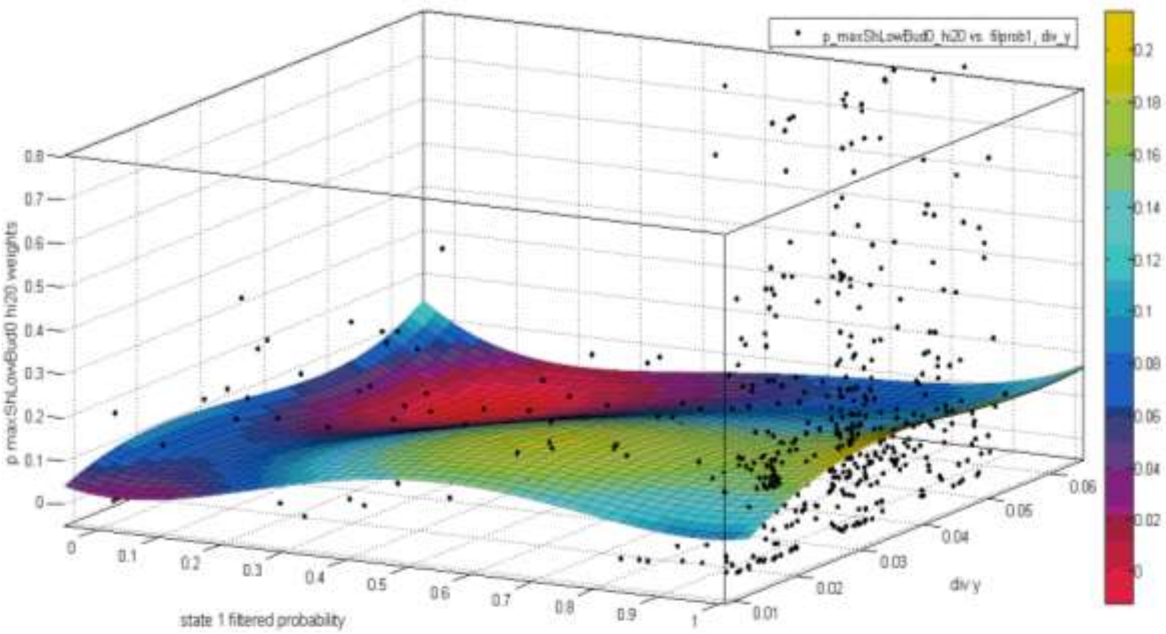


Figure 97 bonds optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with opened lower budget constraint (permit to invest in the

riskless asset) as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

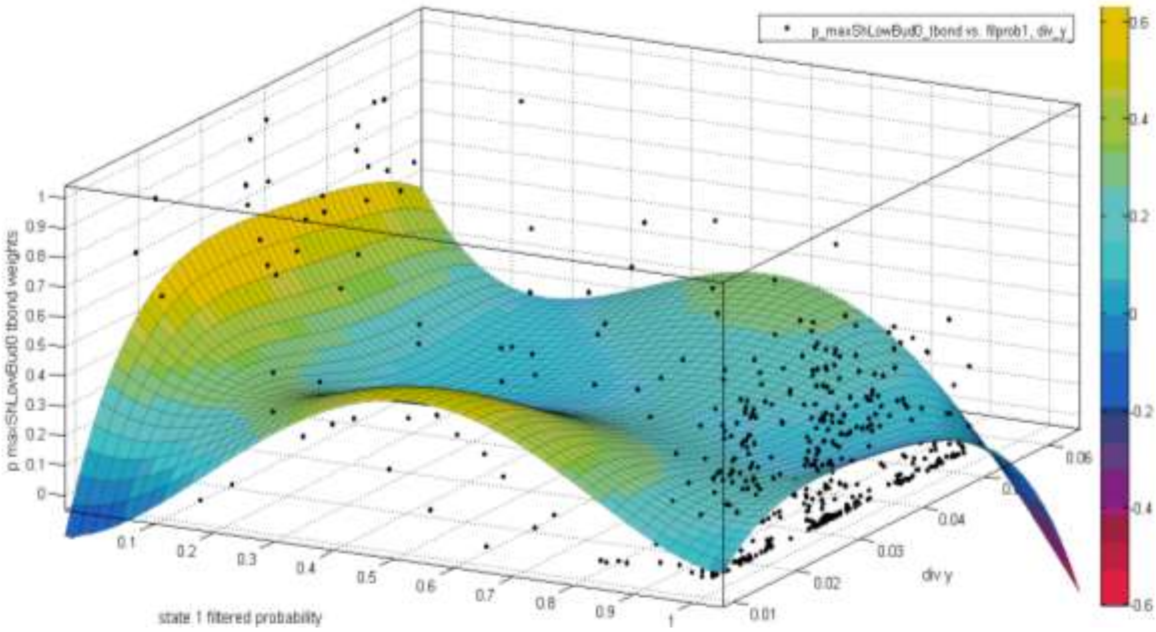


Figure 98 riskless asset optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with opened lower budget constraint (permit to invest in the riskless asset) as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

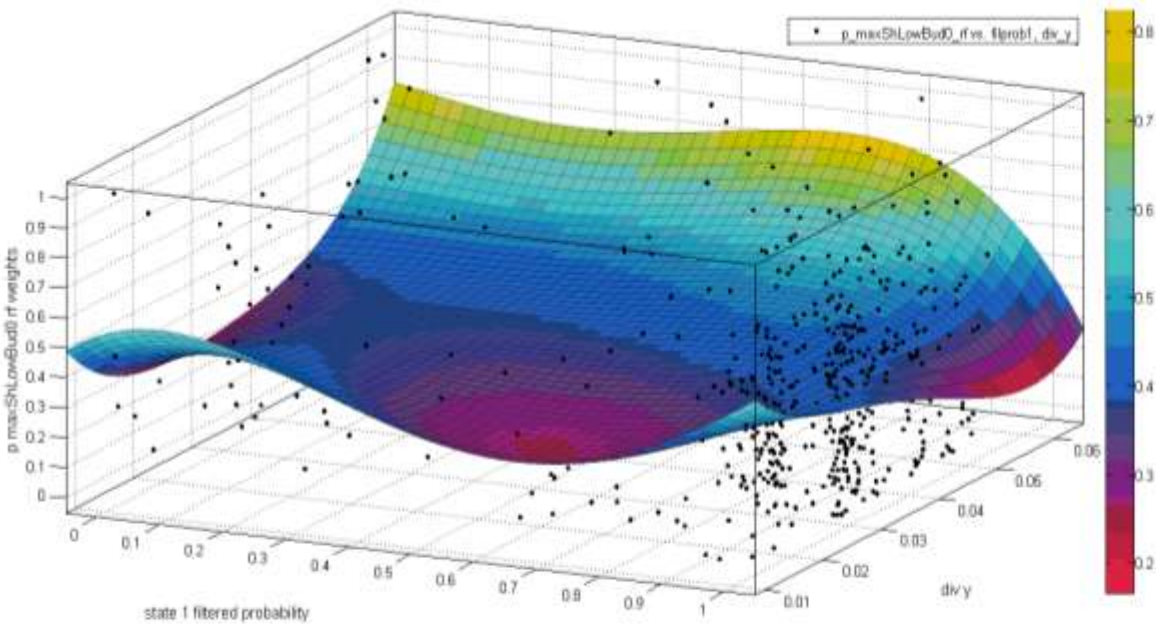


Figure 99 small stocks optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with budget constraint (not permit to invest in the riskless asset) as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

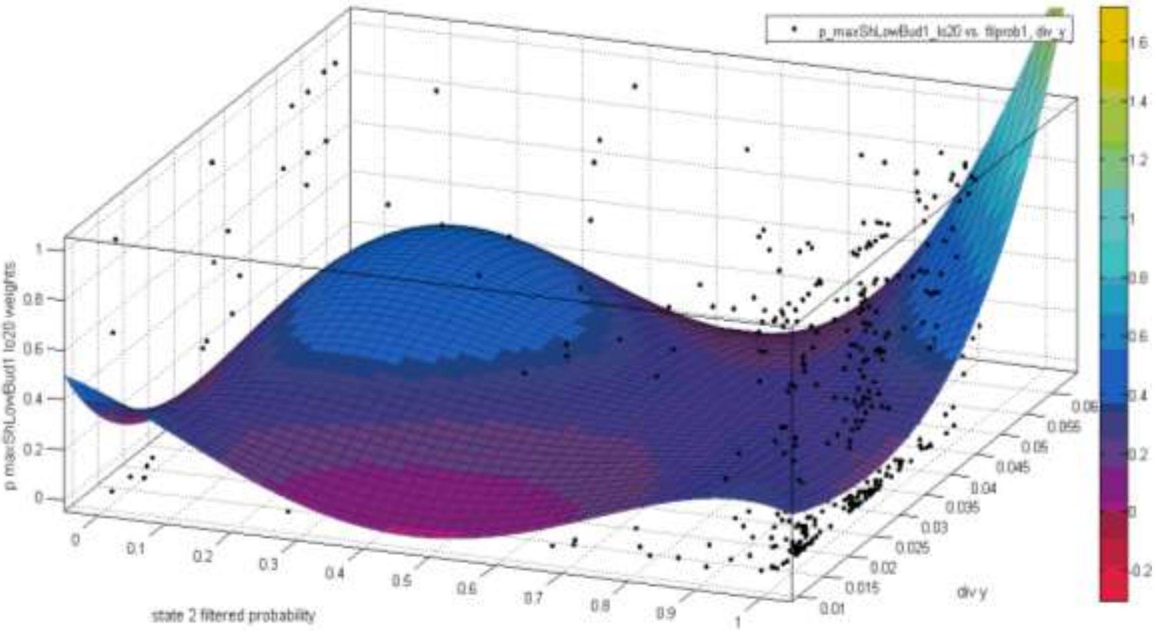


Figure 100 large stocks optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with budget constraint (not permit to invest in the riskless asset) as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

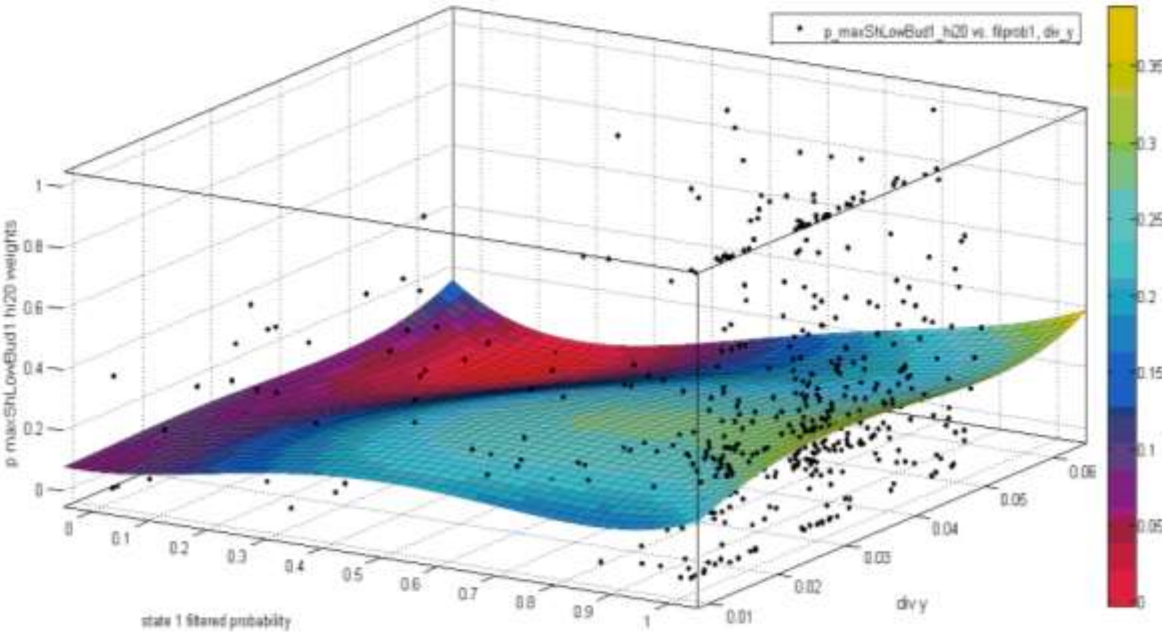


Figure 101 bonds optimal weights (1 = 100%) of the maximum Sharpe ratio portfolios with budget constraint (not permit to invest in the riskless asset) as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

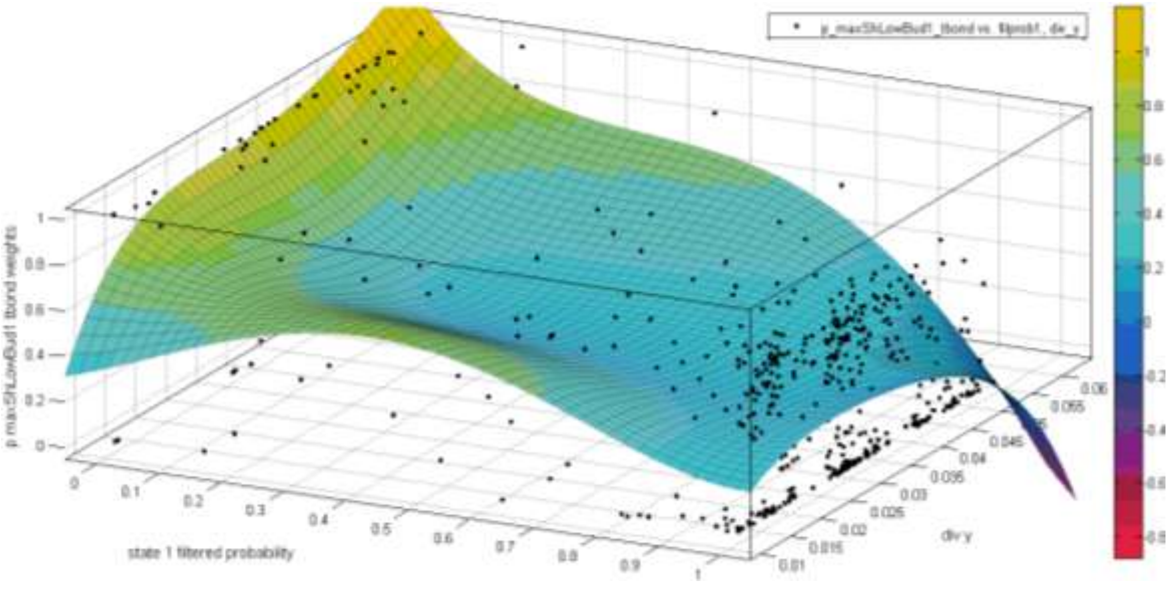


Figure 102 small stocks optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 1 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

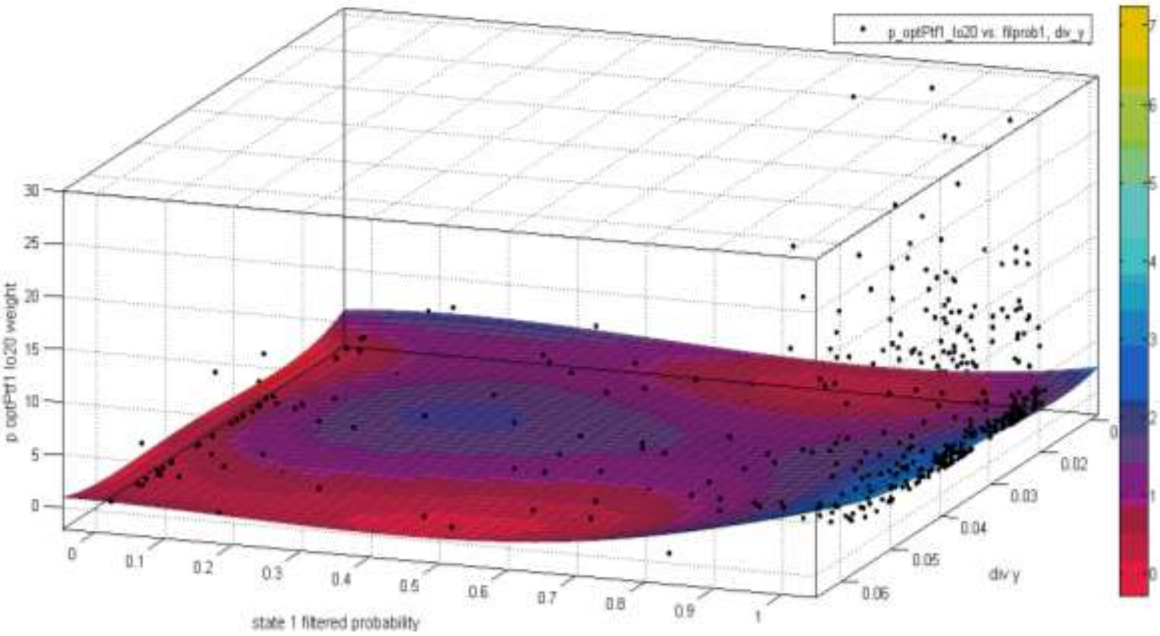


Figure 103 large stocks optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 1 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

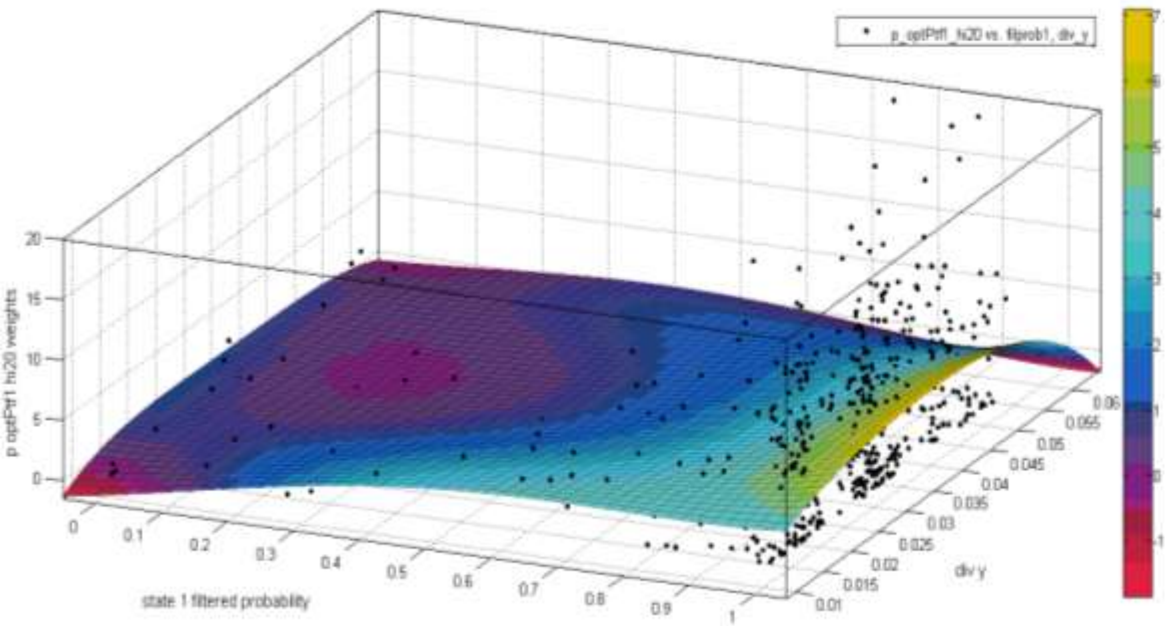


Figure 104 bonds optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 1 as a function of the filtered

probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

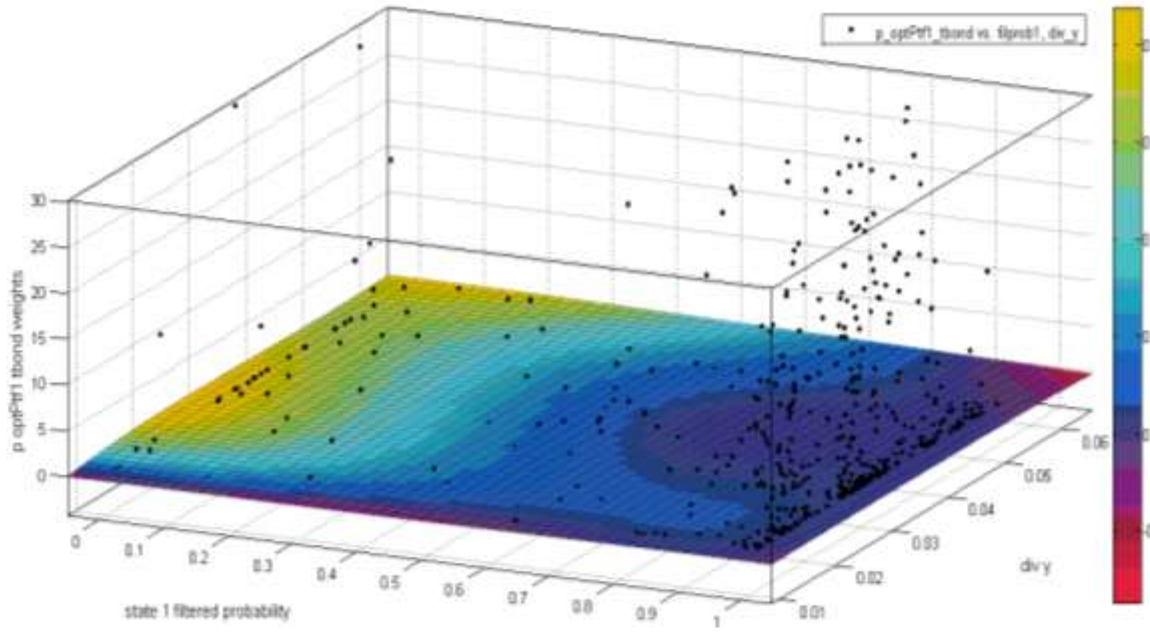


Figure 105 riskless asset optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 1 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

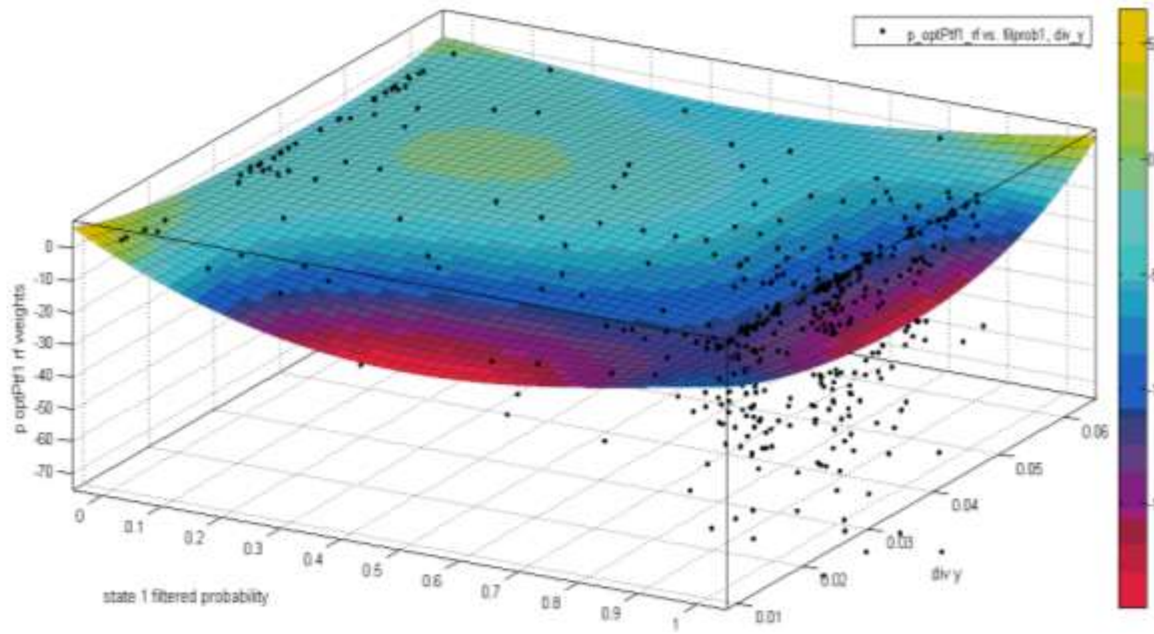


Figure 106 small stocks optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 3 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

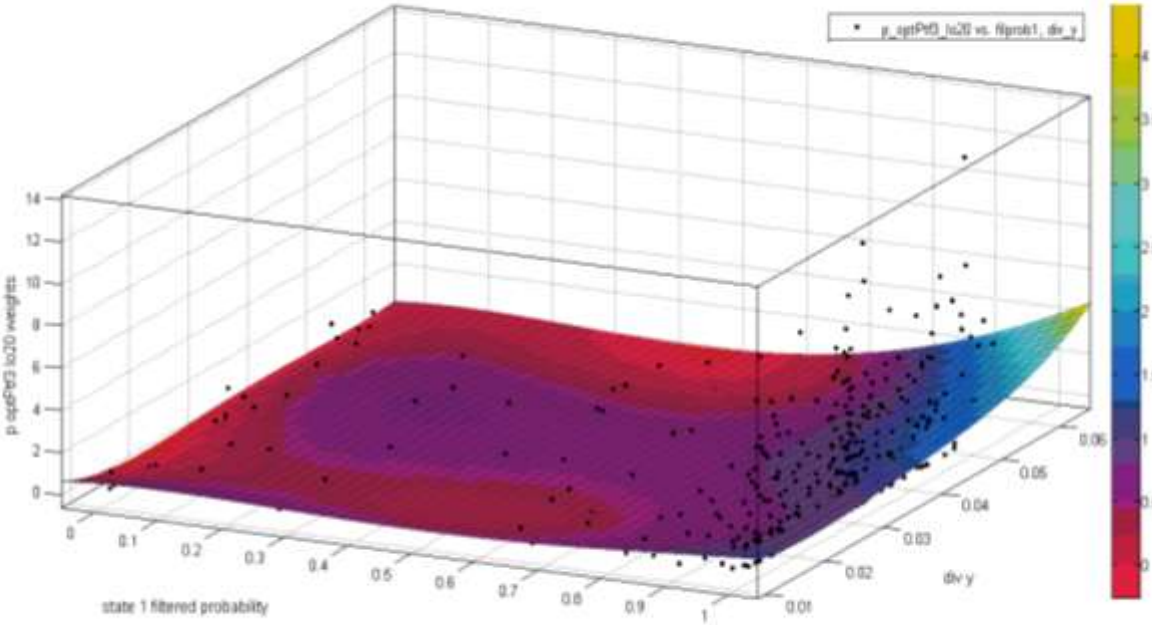


Figure 107 large stocks optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 3 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

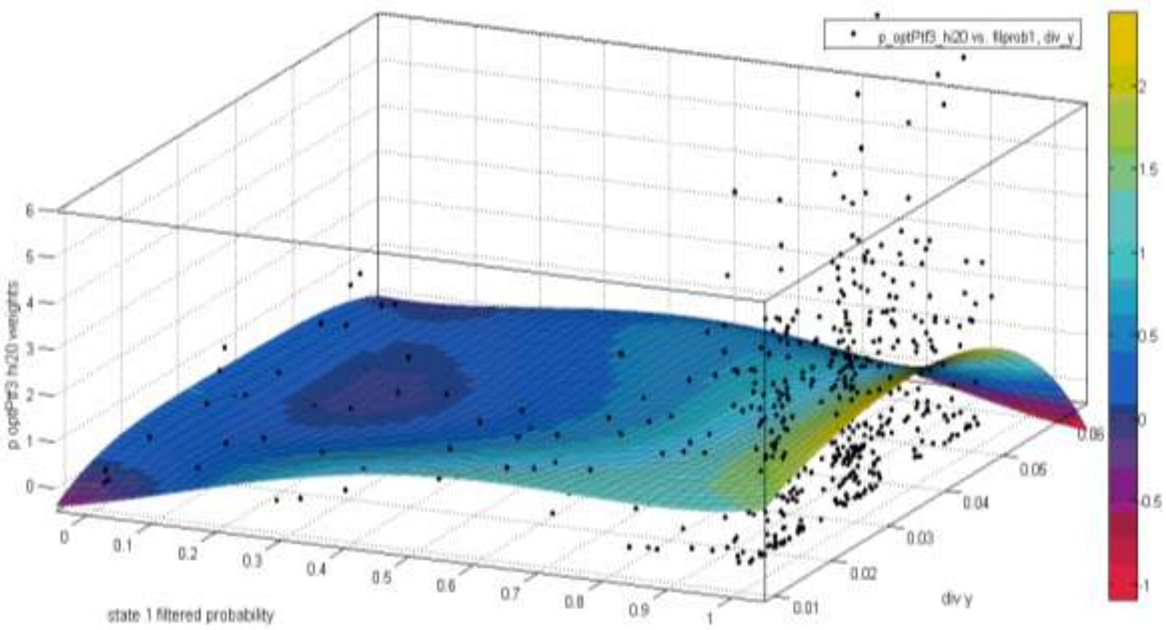


Figure 108 bonds optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 3 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

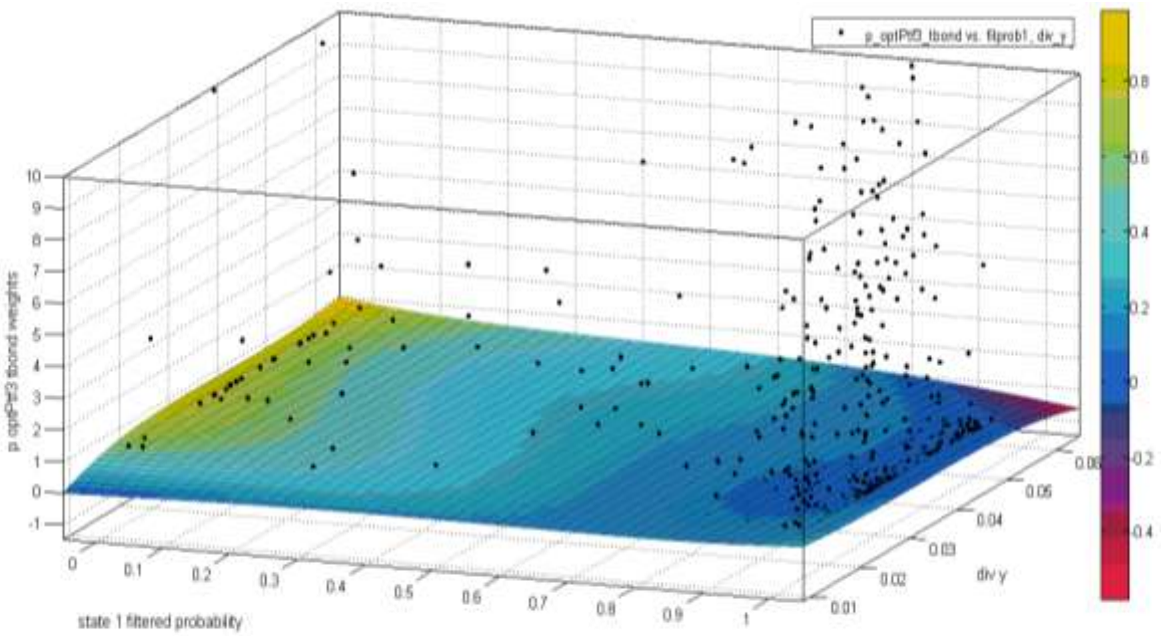


Figure 109 riskless asset optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 3 as a function of the filtered

probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

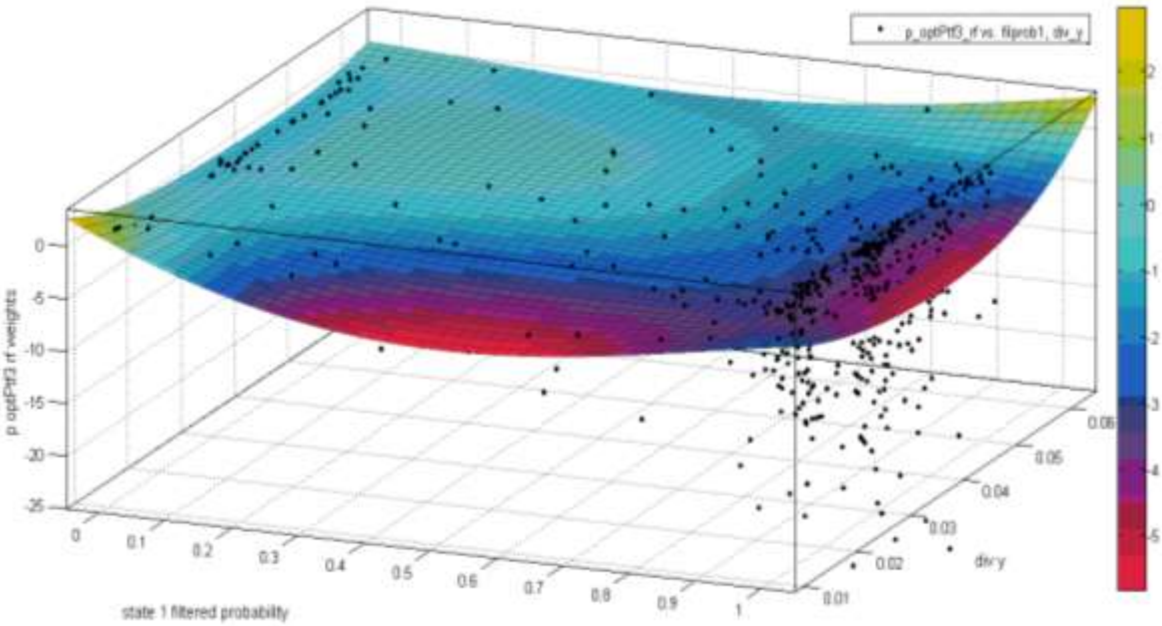


Figure 110 small stocks optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 5 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

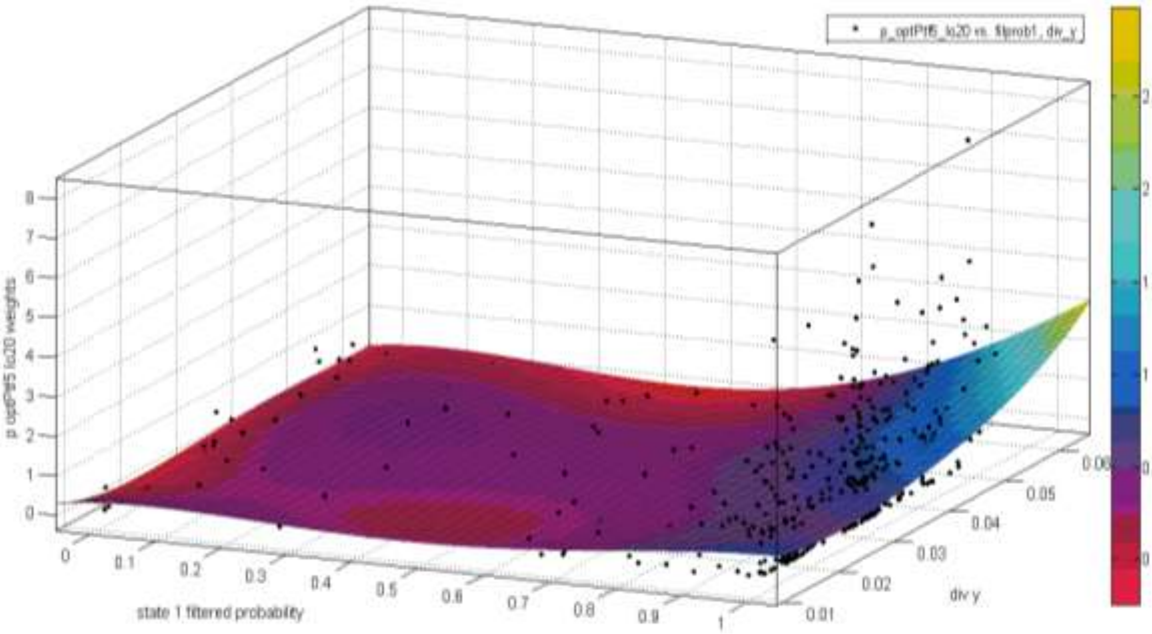


Figure 111 large stocks optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 5 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

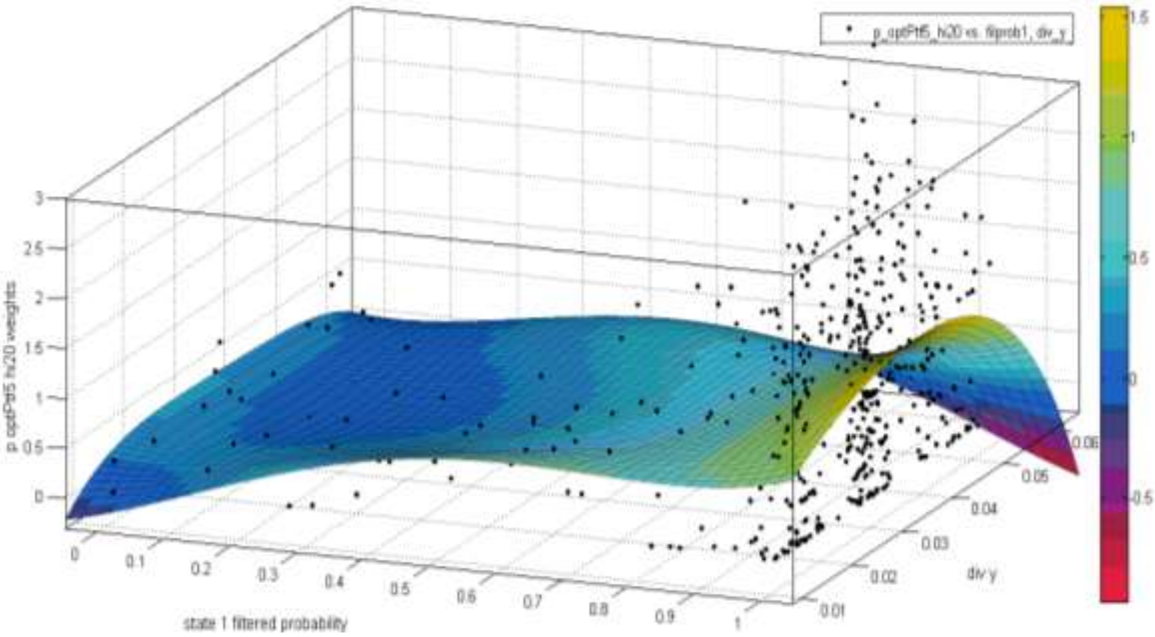


Figure 112 bonds optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 5 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen

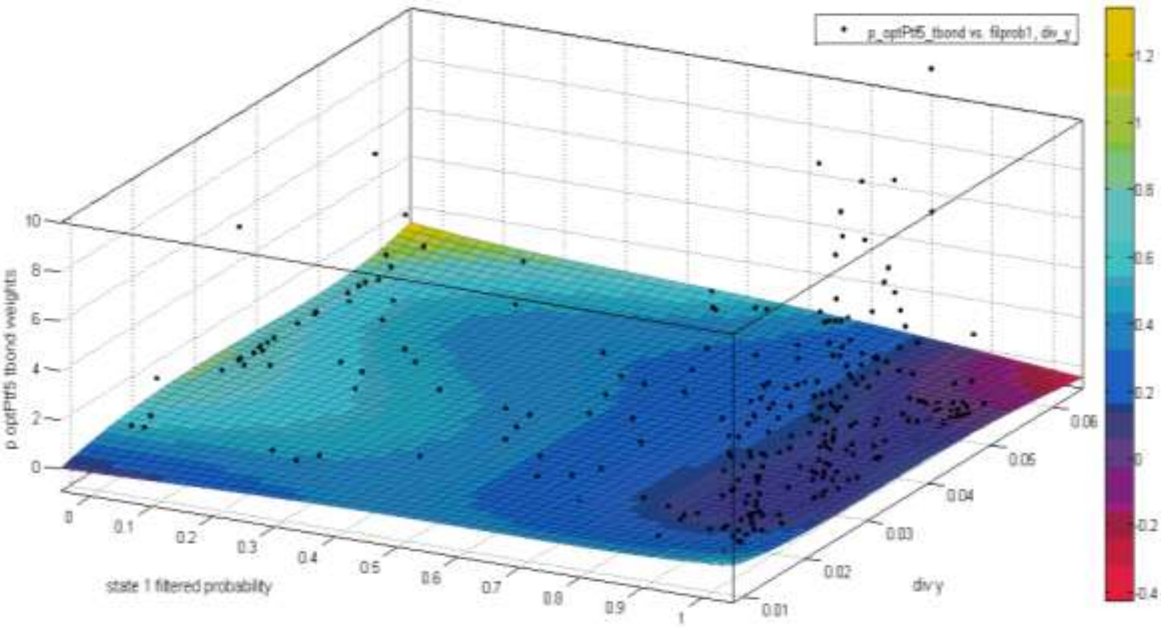
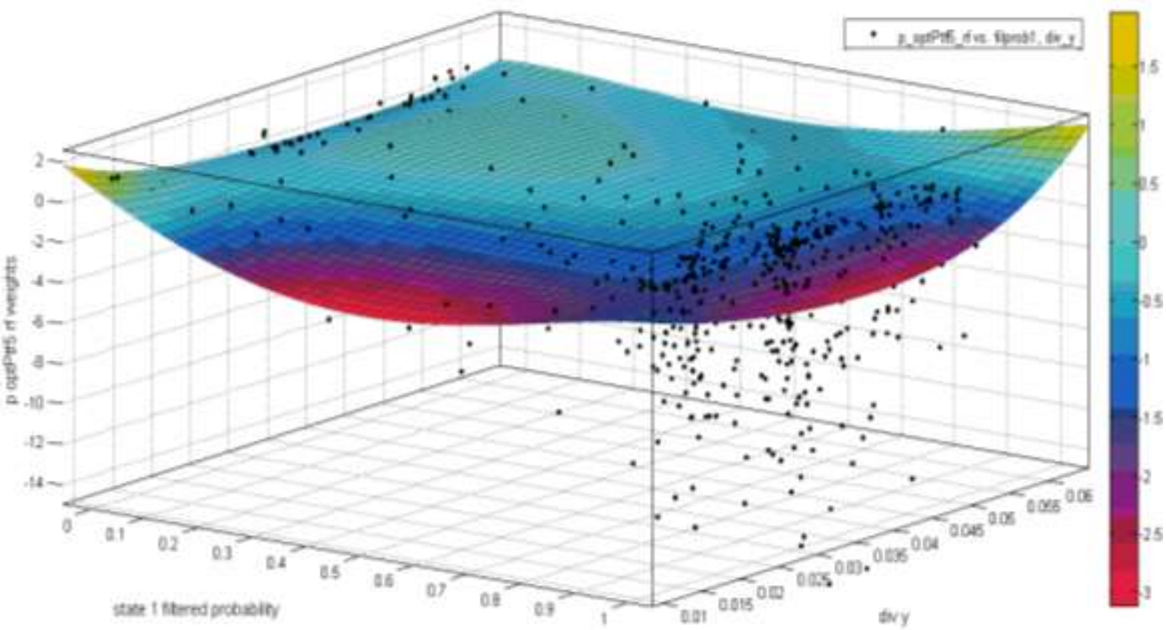


Figure 113 riskless asset optimal weights (1 = 100%) of the portfolios with capital allocation and risk aversion coefficient = 5 as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen



Figures 95 to 113 represent the optimal weights (1 = 100%) of the five different portfolios as a function of the filtered probability of being in state 1, and of the dividend yield at the time the portfolio is chosen. The graphs have been obtained

fitting 4th degrees bisquare robust polynomial surfaces (4 degrees for each input dimensions) using the Matlab Curve Fitting Tool. As it can be seen for the first portfolio, the one that maximizes the expected Sharpe ratio with opened lower budget constraint, the lo20 optimal weights on average grow as a function of state 1 perceived probability especially for high dividend yield values; the same overall trend can be spotted in the hi20 optimal weights surface even though its slope is steeper and less sensitive to the dividend yields; conversely, because of its hedging properties due to low volatility and a high state 2 mean return, the tbond is characterized on average by decreasing optimal weights as a function of state 1 filtered probability. The conclusions that can be drawn from the observation of the surfaces of the portfolio that maximizes the expected Sharpe ratio with budget constraint are pretty similar to the ones just discussed. From the analysis of the last three portfolios, those chosen to maximize the expected utility function given three different risk aversion levels, emerges that on average the hi20 optimal weights always increase as a function of the state 1 filtered probability due to its high mean return in this state, in addition it can be seen that conditionally on the dividend yield the hi20 optimal weights are higher for average dividend yield values while are lower for dividend yield values smaller than its mean. The lo20 optimal weights seem to be always pretty constant, while the tbond optimal weights decrease as function of the perceived state 1 probability because of its attractiveness due to its higher mean and low volatility in state 2.

Table 22 in-sample average portfolio overall weights (1 = 100%)

avg. weights maximum Sharpe ratio portfolio with opened lower budget constraint						
	unconditional		*state 1*		*state 2*	
lo20	0.1944	(0.2522)	0.2011	(0.2499)	0.1566	(0.2633)
hi20	0.2045	(0.2209)	0.2294	(0.2276)	0.0634	(0.0919)
tbond	0.1979	(0.2284)	0.1706	(0.1935)	0.3528	(0.3296)
risk-free asset	0.4032	(0.2603)	0.3989	(0.2547)	0.4272	(0.2908)
avg. weights maximum Sharpe ratio portfolio with budget constraint						
	unconditional		*state 1*		*state 2*	
lo20	0.3153	(0.3566)	0.3243	(0.3514)	0.2640	(0.3834)
hi20	0.3509	(0.3028)	0.3909	(0.3011)	0.1241	(0.1937)
tbond	0.3338	(0.3320)	0.2848	(0.2836)	0.6119	(0.4362)
risk-free asset	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)

avg. optimal weights with capital allocation and risk aversion coefficient = 1						
	unconditional		*state 1*		*state 2*	
lo20	3.8171	(5.9090)	4.3355	(6.2342)	0.8796	(1.5717)
hi20	4.7067	(5.3590)	5.3730	(5.4988)	0.9307	(1.8449)
tbond	6.5800	(11.7046)	7.0450	(12.2000)	3.9450	(7.9150)
risk-free asset	-14.1038	(18.7596)	-15.7535	(19.5087)	-4.7552	(9.3859)
avg. optimal weights with capital allocation and risk aversion coefficient = 3						
	unconditional		*state 1*		*state 2*	
lo20	1.3565	(1.9411)	1.5402	(2.0394)	0.3160	(0.5346)
hi20	1.6341	(1.7505)	1.8653	(1.7839)	0.3239	(0.6183)
tbond	2.2351	(3.8847)	2.3525	(4.0647)	1.5695	(2.5608)
risk-free asset	-4.2257	(6.1148)	-4.7580	(6.3676)	-1.2094	(2.9932)
avg. optimal weights with capital allocation and risk aversion coefficient = 5						
	unconditional		*state 1*		*state 2*	
lo20	0.8576	(1.1625)	0.9745	(1.2168)	0.1953	(0.3271)
hi20	1.0164	(1.0405)	1.1593	(1.0548)	0.2069	(0.3789)
tbond	1.3668	(2.3238)	1.4162	(2.4370)	1.0870	(1.5134)
risk-free asset	-2.2408	(3.5994)	-2.5499	(3.7525)	-0.4892	(1.7382)

Table 22 shows unconditional and state conditional average optimal weights for each one of the five different portfolios, their standard deviations are reported in parentheses.

As it can be seen, for the first two portfolios, the hi20 represents the largest share of the portfolio, disregarding the riskless asset, conditionally on state 1 while tbond does it conditionally on state 2; hi20 and tbond have the highest mean returns in state 1 and state 2 respectively, in addition tbond is also the less volatile asset class conditionally on state 2. As a consequence the investor flies from the high profitable stock markets to the low volatile bond market during state 2, in fact both the lo20 and hi20 shares are higher in state 1 than state 2 while the opposite is true for the tbond. Taking a look at the unconditional weights it can be seen that hi20 again represents the largest share for the second portfolio while the riskless asset does it for the first portfolio, this can be explained by the fact that hi20 has the highest unconditional mean return, however it must be that when it is permit to invest in the riskless asset, the investor ends up to be bettered off by holding a large share of his or her portfolio in the riskless asset in order to rise the expected Sharpe ratio. The last three portfolios, those realized by maximizing the expected utility with lending and borrowing capital allocation line and with three different risk

aversion coefficients, are characterized by average optimal weights far larger than 100%. The optimal weights of these three portfolios are characterized by the same rank regardless of whether the optimal weights are considered unconditionally or conditionally on the state. The tbond always represents the largest portfolio share followed by the hi20 and lo20 respectively; the riskless asset constantly represents important short weights given the wide use of borrowing. The rank just discussed might be explained by the fact that even though the tbond has the lowest mean return it is the less volatile asset class followed indeed by the hi20 and at the end the lo20, in addition hi20 has a mean returns higher than lo20 in state 1 which occurs most of the time, contributing to show off investor the hi20 returns more desirable than the lo20. Taking a look at Table 22 it can be also noticed that the average borrowing is a decreasing function of the risk aversion and that it is on average higher during state 1 due to the low volatility nature of this regime that favors higher market exposures.

Table 23 in-sample average portfolio risky weights (1 = 100%)

avg. optimal risky weights with capital allocation and risk aversion coefficient = 1						
	unconditional		*state 1*		*state 2*	
lo20	0.3519	(0.3886)	0.3654	(0.3856)	0.2753	(0.3985)
hi20	0.3424	(0.3253)	0.3859	(0.3273)	0.0959	(0.1649)
tbond	0.2612	(0.3264)	0.2073	(0.2631)	0.5670	(0.4583)
avg. optimal risky weights with capital allocation and risk aversion coefficient = 3						
	unconditional		*state 1*		*state 2*	
lo20	0.3495	(0.3856)	0.3634	(0.3832)	0.2703	(0.3920)
hi20	0.3435	(0.3235)	0.3866	(0.3256)	0.0989	(0.1636)
tbond	0.2626	(0.3273)	0.2086	(0.2646)	0.5691	(0.4576)
avg. optimal risky weights with capital allocation and risk aversion coefficient = 5						
	unconditional		*state 1*		*state 2*	
lo20	0.3454	(0.3821)	0.3602	(0.3802)	0.2618	(0.3844)
hi20	0.3452	(0.3217)	0.3880	(0.3236)	0.1027	(0.1664)
tbond	0.2649	(0.3270)	0.2104	(0.2644)	0.5738	(0.4546)

Table 23 again shows that the investor during state 1 prefers to hold a risky portfolio that mainly consists of hi20 while during state 2 a portfolio that mainly consists of tbond; these risky portfolio are then leveraged by a borrowing factor which in turn is influenced by the lending and borrowing rates, the risk aversion coefficient and the market returns. From a comparison of Table 22 and 23 emerges some conflicting information; although the overall tbond optimal weights are the largest, from Table 22 emerges that the share of this asset class in then average risky portfolios is not the largest, this might be explained by the fact that the investor might have largely resort to leverage when he or she has invested a prevailing part of the portfolio in the tbond asset class, this is the case during state 2, in fact in this state tbond weights dominate the other by a great extent, in addition the same large resort to borrowing might have been occurred also in certain periods classified as state 1 when tbond represented a large portfolio share, as a result the contribution of the leverage at certain periods caused the relative weight of the tbond asset class to grow, compared with the other asset classes weights, making weights on average the largest in each one of the three maximum expected utility portfolio.

Table 24 in-sample portfolios expected moments

maximum Sharpe ratio portfolio with opened lower budget constraint						
	unconditional		*state 1*		*state 2*	
avg. expected return	0.0073	(0.0071)	0.0072	(0.0071)	0.0075	(0.0068)
avg. expected standard deviaton	0.0167	(0.0123)	0.0155	(0.0106)	0.0234	(0.0179)
avg. Sharpe ratio	0.4668	(0.3309)	0.4877	(0.3374)	0.3484	(0.2631)
maximum Sharpe ratio portfolio with budget constraint						
	unconditional		*state 1*		*state 2*	
avg. expected return	0.0113	(0.0093)	0.0112	(0.0094)	0.0117	(0.0088)
avg. expected standard deviaton	0.0280	(0.0138)	0.0258	(0.0112)	0.0407	(0.0195)
avg. Sharpe ratio	0.4634	(0.3363)	0.4846	(0.3426)	0.3434	(0.2706)
optimal weights with capital allocation and risk aversion coefficient = 1						
	unconditional		*state 1*		*state 2*	
avg. expected return	0.2056	(0.3995)	0.2254	(0.4224)	0.0933	(0.1965)

avg. expected standard deviaton	0.3144	(0.3199)	0.3381	(0.3269)	0.1799	(0.2370)
avg. Sharpe ratio	0.4005	(0.2982)	0.4153	(0.3050)	0.3167	(0.2409)

optimal weights with capital allocation and risk aversion coefficient = 3

	unconditional		*state 1*		*state 2*	
avg. expected return	0.0719	(0.1325)	0.0784	(0.1401)	0.0351	(0.0649)
avg. expected standard deviaton	0.1121	(0.1009)	0.1195	(0.1033)	0.0702	(0.0736)
avg. Sharpe ratio	0.4085	(0.2970)	0.4236	(0.3034)	0.3229	(0.2416)

optimal weights with capital allocation and risk aversion coefficient = 5

	unconditional		*state 1*		*state 2*	
avg. expected return	0.0450	(0.0792)	0.0489	(0.0838)	0.0233	(0.0388)
avg. expected standard deviaton	0.0710	(0.0581)	0.0752	(0.0596)	0.0474	(0.0421)
avg. Sharpe ratio	0.4154	(0.2969)	0.4310	(0.3030)	0.3275	(0.2434)

Table 24 shows, unconditional and state conditional, average expected returns, average expected standard deviations and average Sharpe ratios of the five different portfolios. The average expected returns are the single period portfolios expected returns, function of the underlying asset classes expected returns and portfolio weights, averaged over the whole in-sample period. Coherently with the average expected returns, the average expected standard deviations are the single period standard deviations of the underlying asset class in which the portfolios invest, weighted by the portfolios weights and corrected for the regime switching dynamic, and averaged over the whole in-sample period; the Sharpe ratios again are the single period ratios, function of the portfolio expected returns and expected standard deviations, averaged over the whole period. It can be seen, as largely expected, that for the last three portfolios the average expected return during state 1 is higher than during state 2, surprisingly the opposite is true for the first two portfolios even thought the differences are very small. Surprisingly, the average expected portfolio standard deviations of the last three portfolios are higher in state 1 than in state 2 and not the opposite, as widely expected and indeed happened for the first two portfolios where the state 2 portfolio average expected deviations are almost twice as large as in state 1. This might be explained by the wide optimal weights oscillations due to the large resort to

borrowing, this conclusion can be confirmed by the fact that, both the unconditional and the state conditional, average portfolios expected standard deviations decrease as a function of the increase of the risk aversion coefficient which in turn means that the higher the risk aversion coefficient, the lower the average resort to borrowing, the lower the average portfolio expected standard deviations. Taking a look at the unconditional average portfolio expected returns and standard deviations it can be noticed that, as largely expected, the higher the risk aversion coefficient the less profitable and less risky in terms of standard deviations the resulting portfolio. Interestingly giving the investor the chance to hold some riskless asset when it maximize the Sharpe ratio leads to a lower average portfolio expected returns and average portfolio expected standard deviations that contributes to the realization of a higher expected Sharpe ratio. The last three portfolios are characterized by an increasing unconditional average expected Sharpe ratio as a function of the amount borrowed to invest in the risky assets. Overall it is clearly noticed that the average state 1 expected Sharpe ratios greatly exceed the correspondent state 2 measures which means that the estimated regimes have a clear economic interpretation observable also from the optimal weights estimated for the five different portfolios.

Figure 114 maximum Sharpe ratio portfolios (with and without budget constraint) expected returns

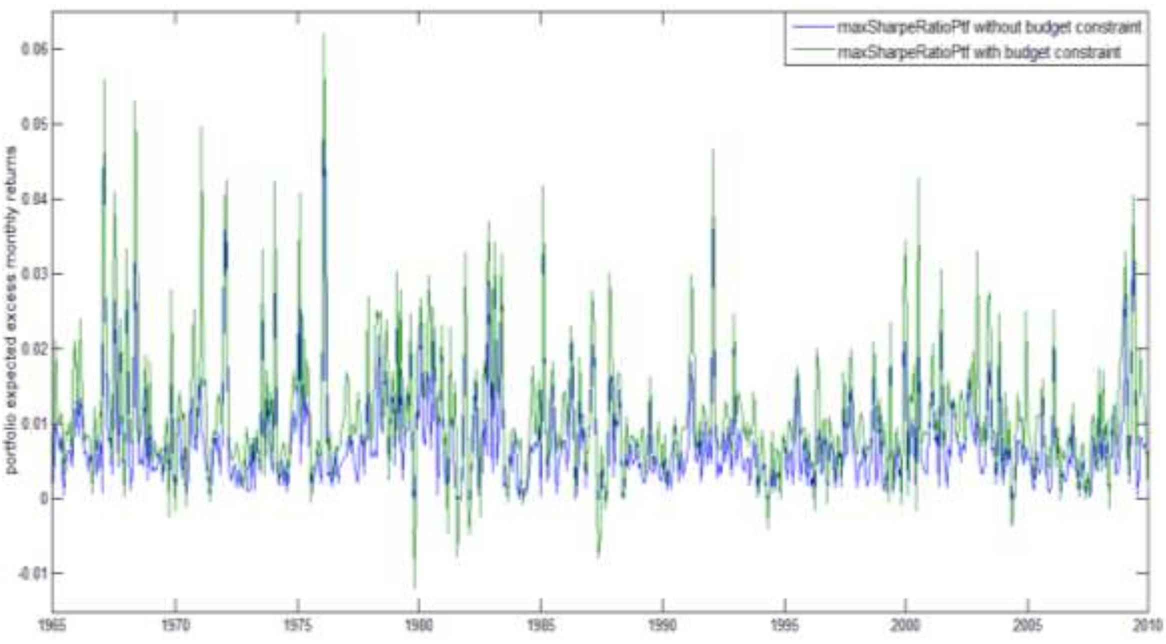
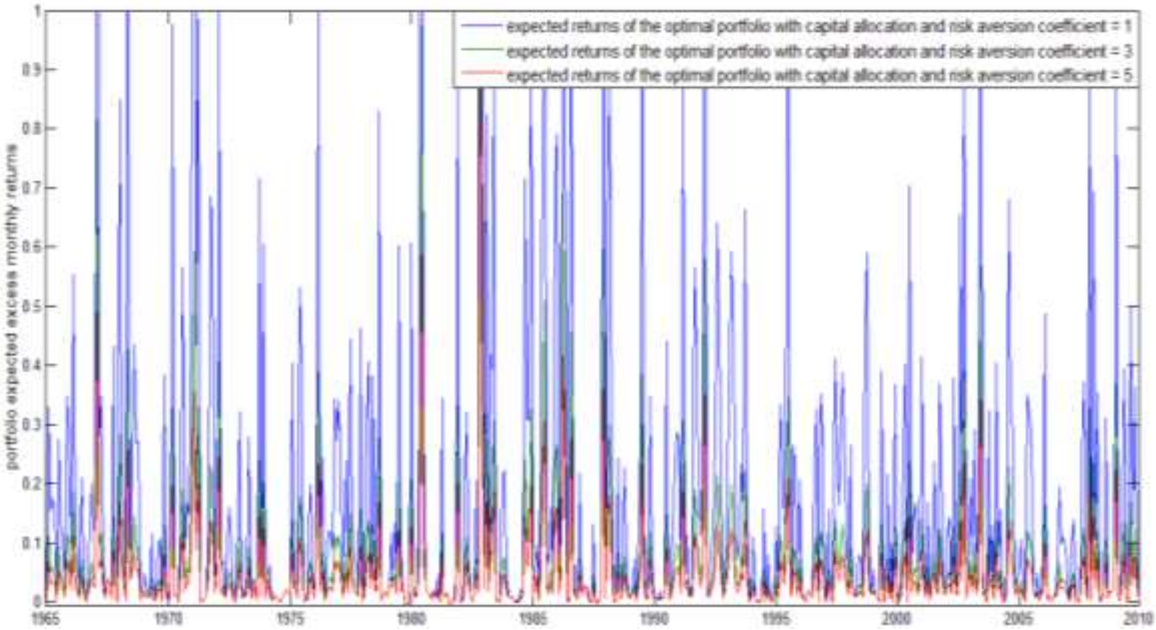


Figure 115 expected returns of the maximum expected utility portfolios (capped at 1 = 100%)



In this section of the subparagraph the in-sample portfolio realized results are shown.

Table 25 in-sample realized portfolios statistics

maximum Sharpe ratio portfolio with opened lower budget constraint

	unconditional	*state 1*	*state 2*
average realized return	0.0070	0.0073	0.0052
realized returns standard deviaton	0.0264	0.0210	0.0463
realized returns Sharpe ratio	0.2650	0.3469	0.1126

maximum Sharpe ratio portfolio with budget constraint

	unconditional	*state 1*	*state 2*
average realized return	0.0102	0.0112	0.0041
realized returns standard deviaton	0.0407	0.0310	0.0751
realized returns Sharpe ratio	0.2492	0.3620	0.0544

optimal weights with capital allocation and risk aversion coefficient = 1

	unconditional	*state 1*	*state 2*
average realized return	0.1979	0.2193	0.0769
realized returns standard deviaton	0.5934	0.6232	0.3628
realized returns Sharpe ratio	0.3336	0.3519	0.2119

optimal weights with capital allocation and risk aversion coefficient = 3

	unconditional	*state 1*	*state 2*

average realized return	0.0690	0.0763	0.0278
realized returns standard deviaton	0.1985	0.2082	0.1239
realized returns Sharpe ratio	0.3477	0.3665	0.2245

optimal weights with capital allocation and risk aversion coefficient = 5

	unconditional	*state 1*	*state 2*
average realized return	0.0432	0.0476	0.0186
realized returns standard deviaton	0.1200	0.1257	0.0767
realized returns Sharpe ratio	0.3600	0.3784	0.2421

As it can be seen, on average, the portfolios realized returns during state 1 overperform those realized during state 2, as largely expected. Similarly to what has already been discussed regarding Table 24, the higher realized portfolio standard deviations during state 1 might be explained by the wide optimal weights oscillations due to the large resort to borrowing, this conclusion can be confirmed by the fact that, both the unconditional that the state conditional, average portfolio expected standard deviations decrease as a function the increase of the risk aversion coefficient which in turn means that the higher the risk aversion coefficient, the lower the average resort to borrowing, thus the lower the average portfolio expected standard deviations. From Table 24 it can be seen that the realized Sharpe ratio increases as a function of the risk aversion coefficient reaching a peak of 0.36 when the former is equal to 5. Again, interestingly, giving the investor the chance to hold some riskless asset when it maximize the Sharpe ratio leads to a lower average portfolio expected returns and average portfolio expected standard deviations that contributes to the realization of a higher expected Sharpe ratio. Overall it can be deduced that the conclusions obtained from the observation of the expected portfolio returns are here substantially reproduced.

Figure 116 in-sample maximum Sharpe ratio portfolio with opened lower budget constraint realized returns

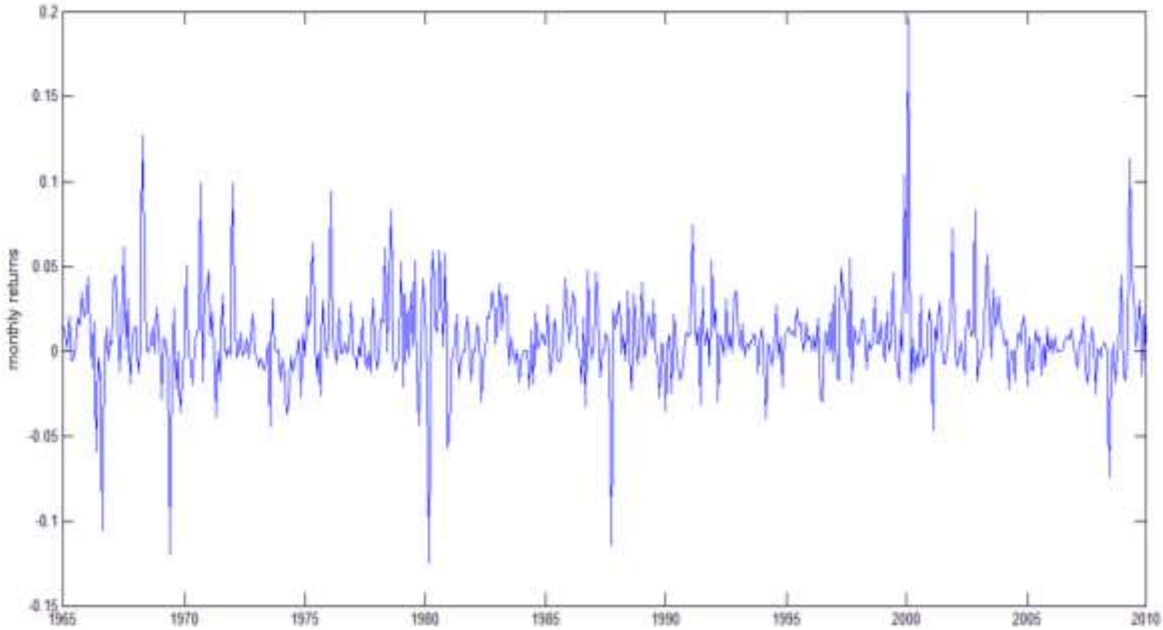


Figure 117 in-sample maximum Sharpe ratio portfolio with opened lower budget constraint cumulative realized returns

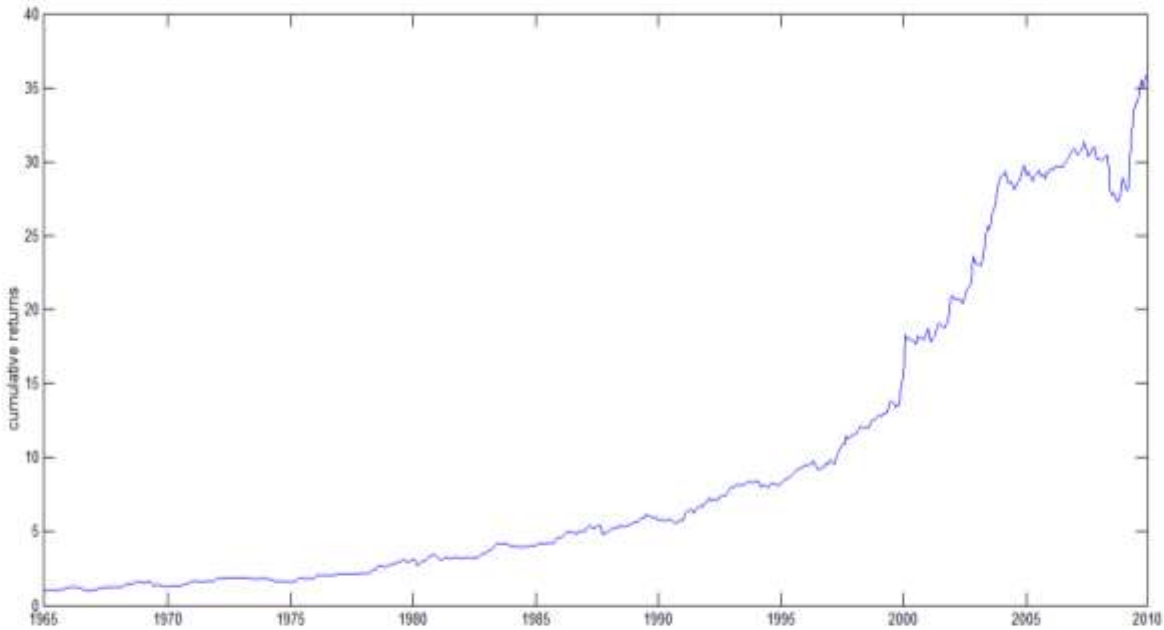


Figure 118 in-sample maximum Sharpe ratio portfolio with budget constraint realized returns

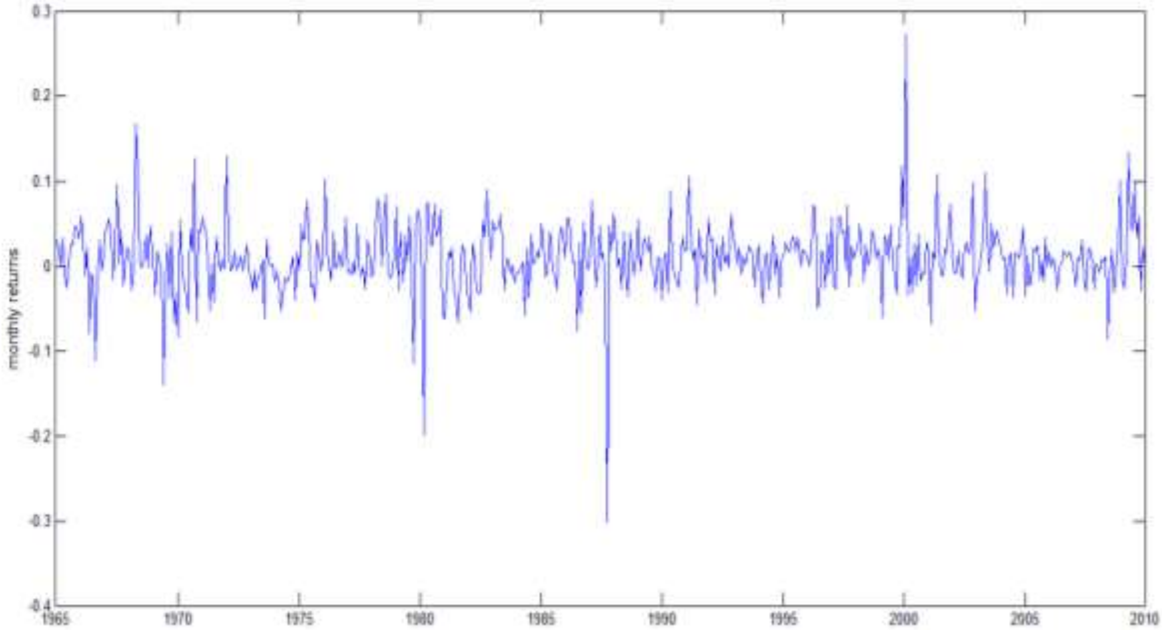


Figure 119 in-sample maximum Sharpe ratio portfolio with budget constraint realized cumulative returns

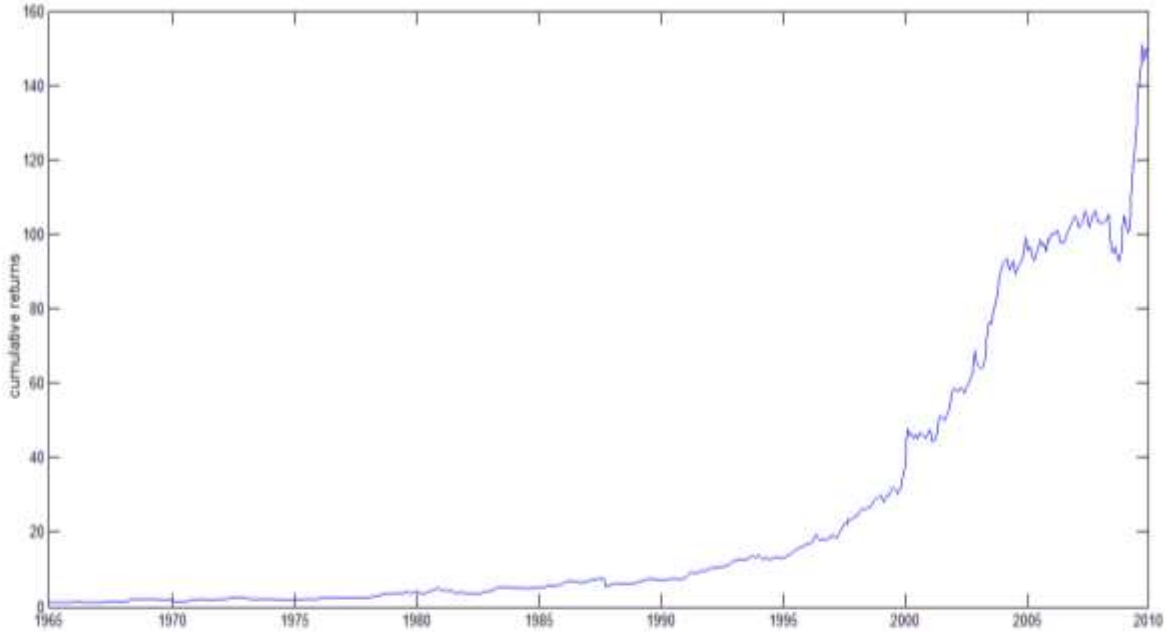


Figure 120 in-sample optimal portfolio with capital allocation and risk aversion coefficient = 1 realized returns

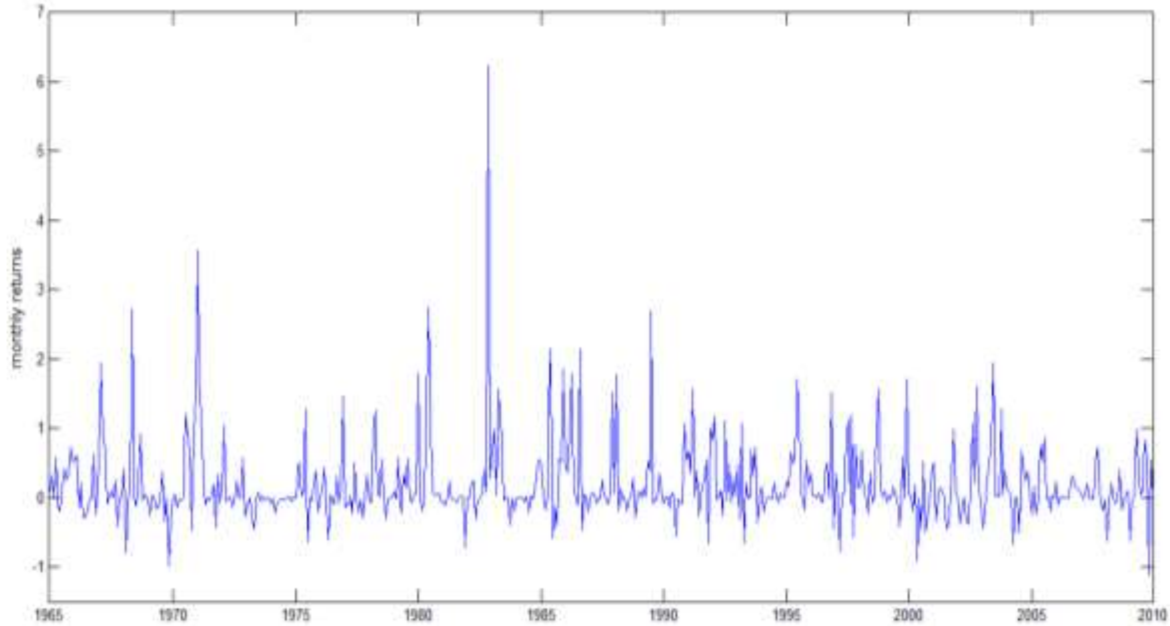


Figure 121 in-sample optimal portfolio with capital allocation and risk aversion coefficient = 1 cumulative realized returns

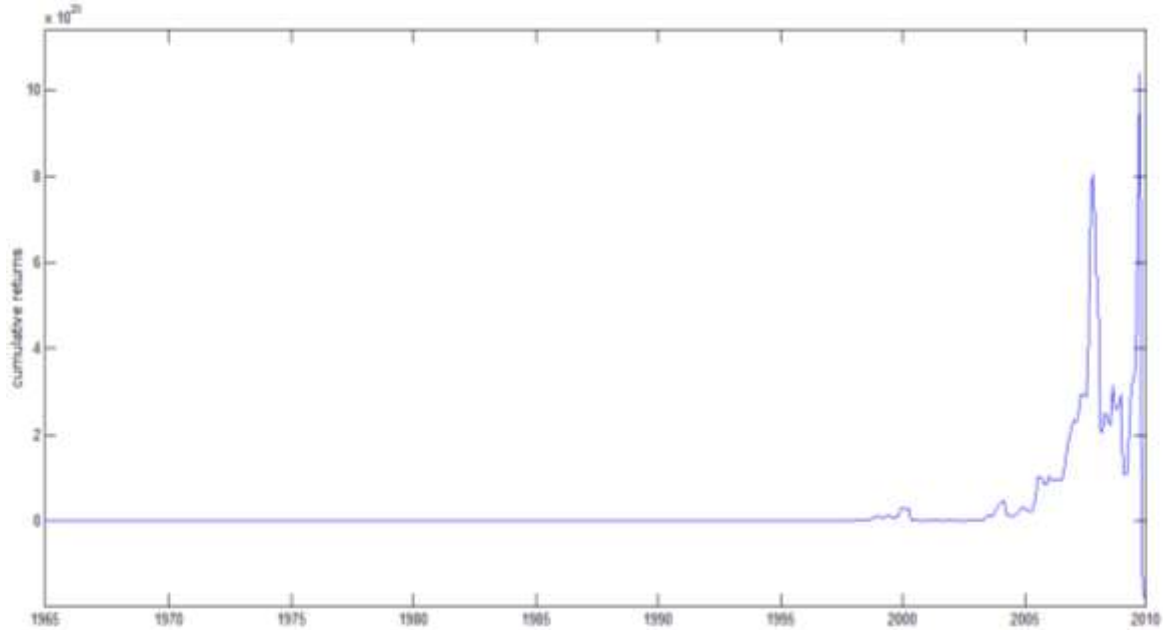


Figure 122 in-sample optimal portfolio with capital allocation and risk aversion coefficient = 3 realized returns

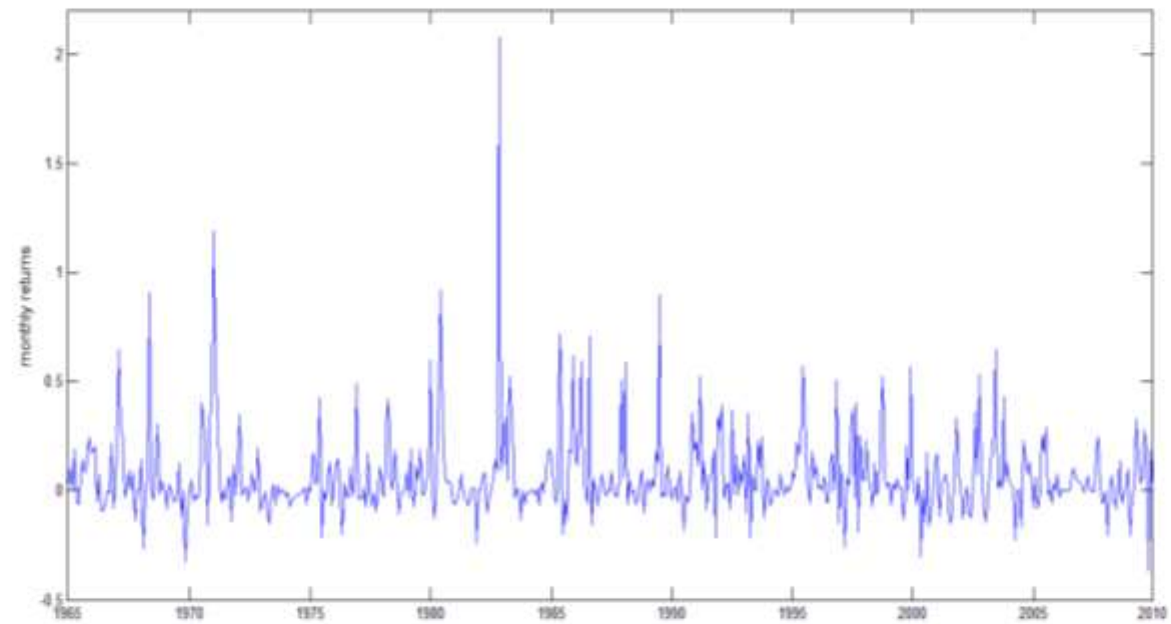


Figure 123 in-sample optimal portfolio with capital allocation and risk aversion coefficient = 3 cumulative realized returns

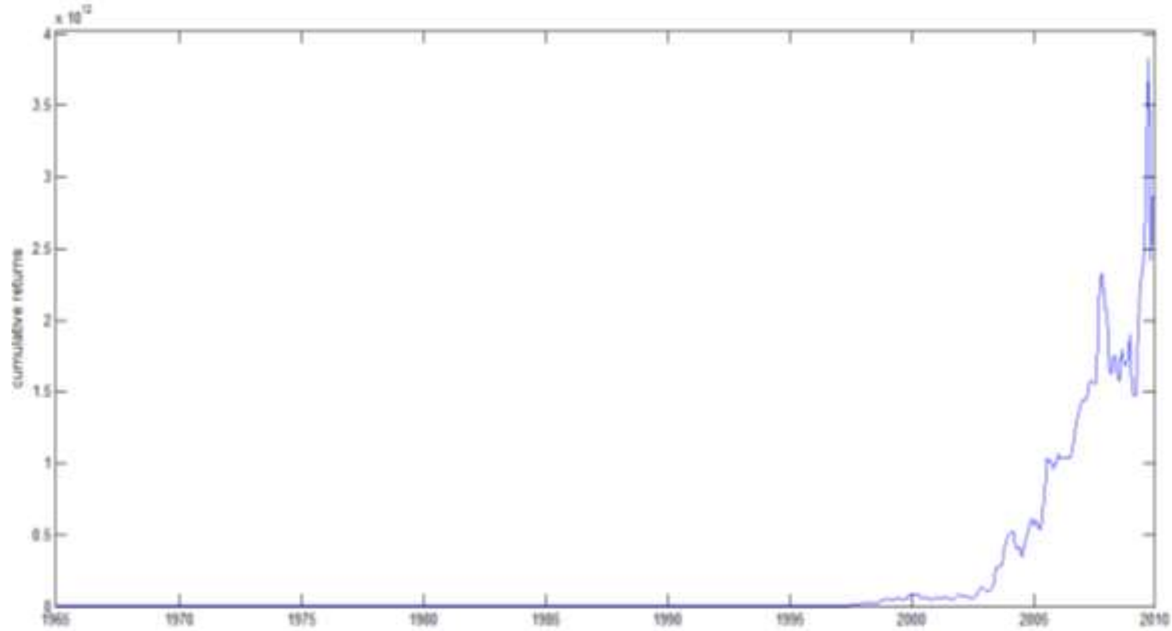


Figure 124 in-sample optimal portfolio with capital allocation and risk aversion coefficient = 5 realized returns

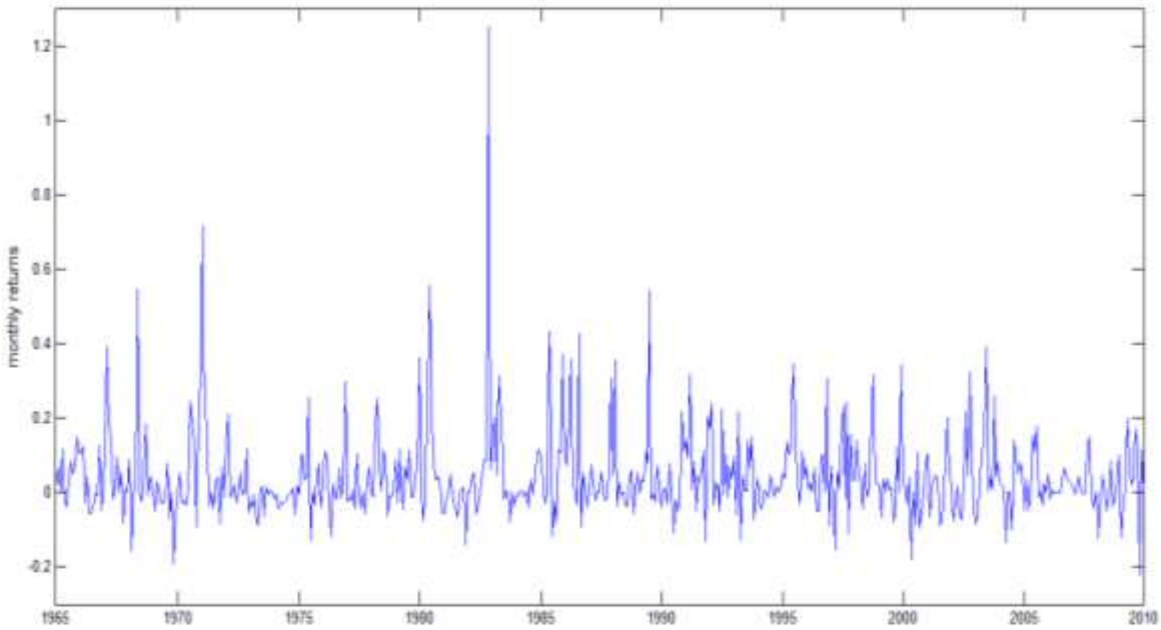
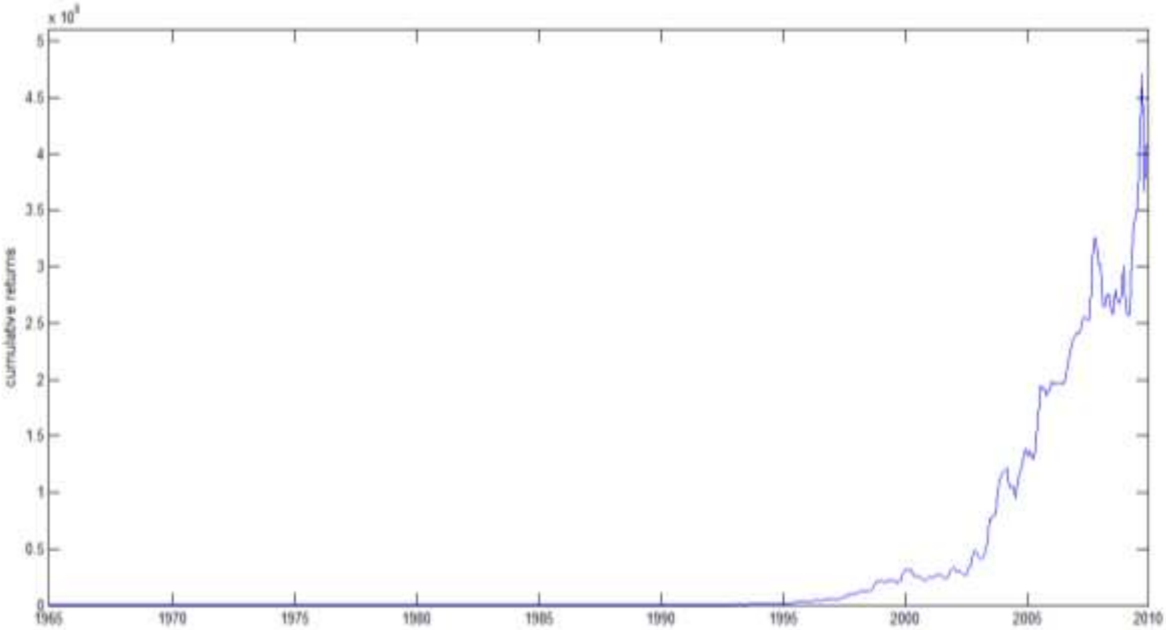


Figure 125 in-sample optimal portfolio with capital allocation and risk aversion coefficient = 5 cumulative realized returns



4.2.2 *In-sample asset allocation exercise based on a VAR(1) model*

In this subparagraph the in-sample portfolio construction results using the VAR(1) are shown. The same five in-sample recursive optimal portfolios as in the MSVAR(2,1) case have been built.

Figure 126 lo20 and its VAR(1) predicted conditional mean

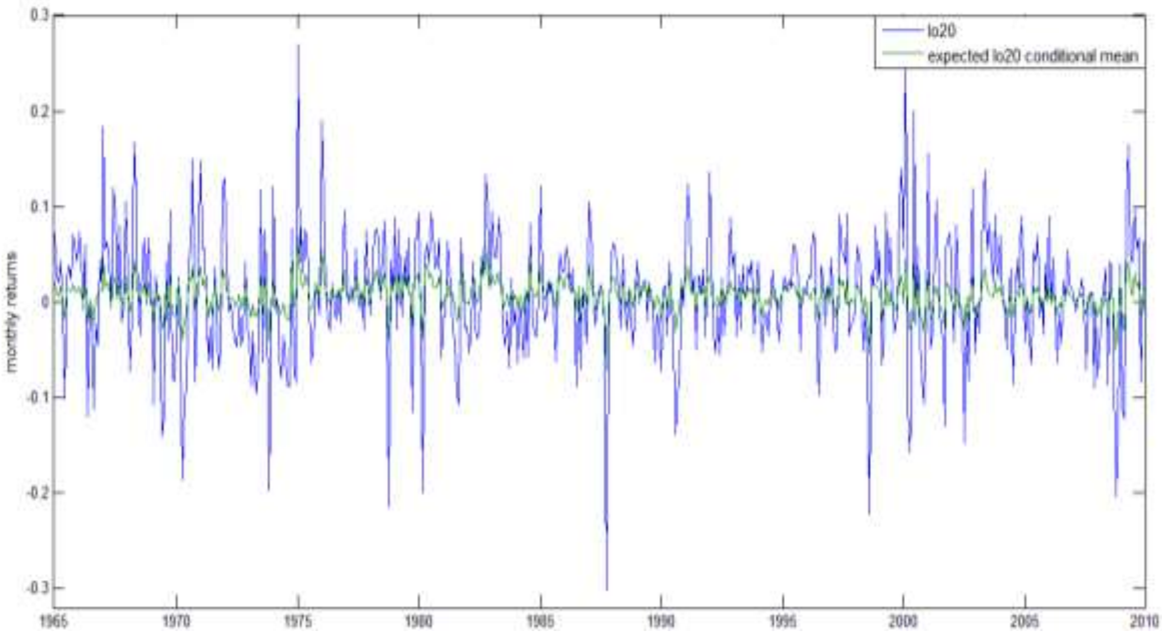


Figure 127 hi20 and its VAR(1) predicted conditional mean

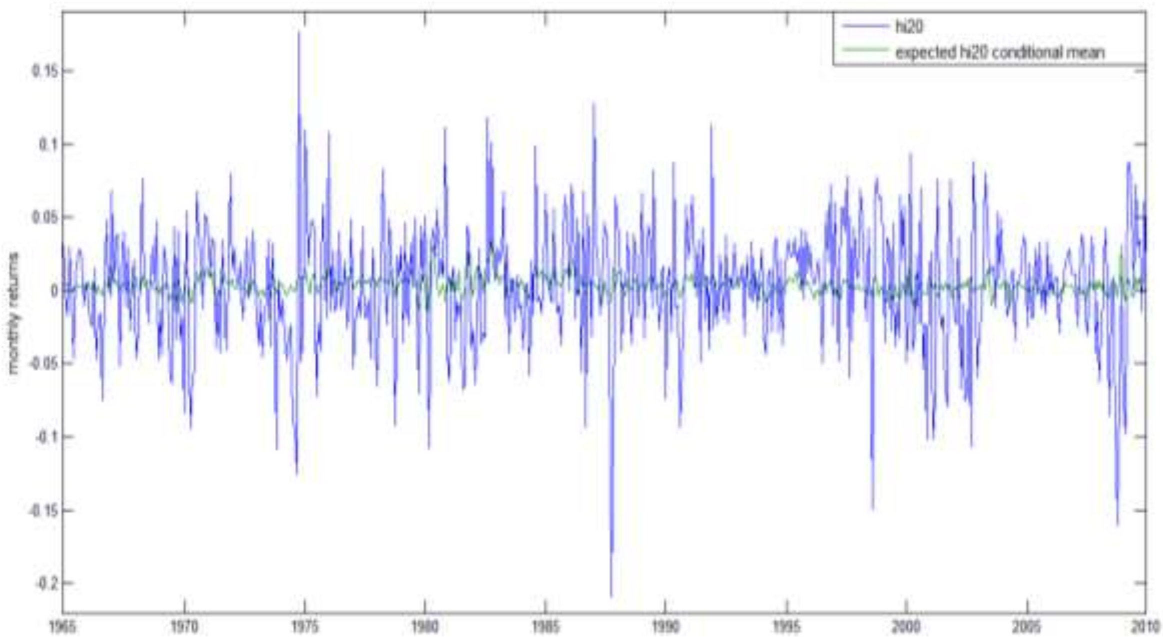


Figure 128 tbond and its VAR(1) predicted conditional mean

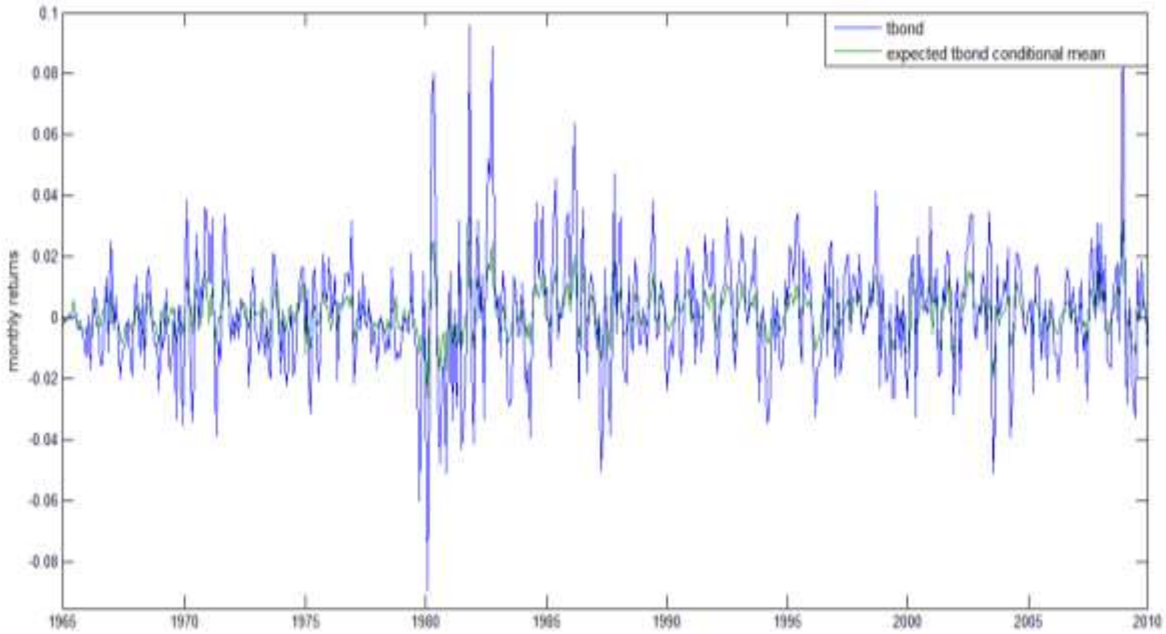


Figure 129 VAR(1) maximum Sharpe ratio portfolios weights (1 = 100%) with opened lower budget constraint (permit to invest in the riskless asset)

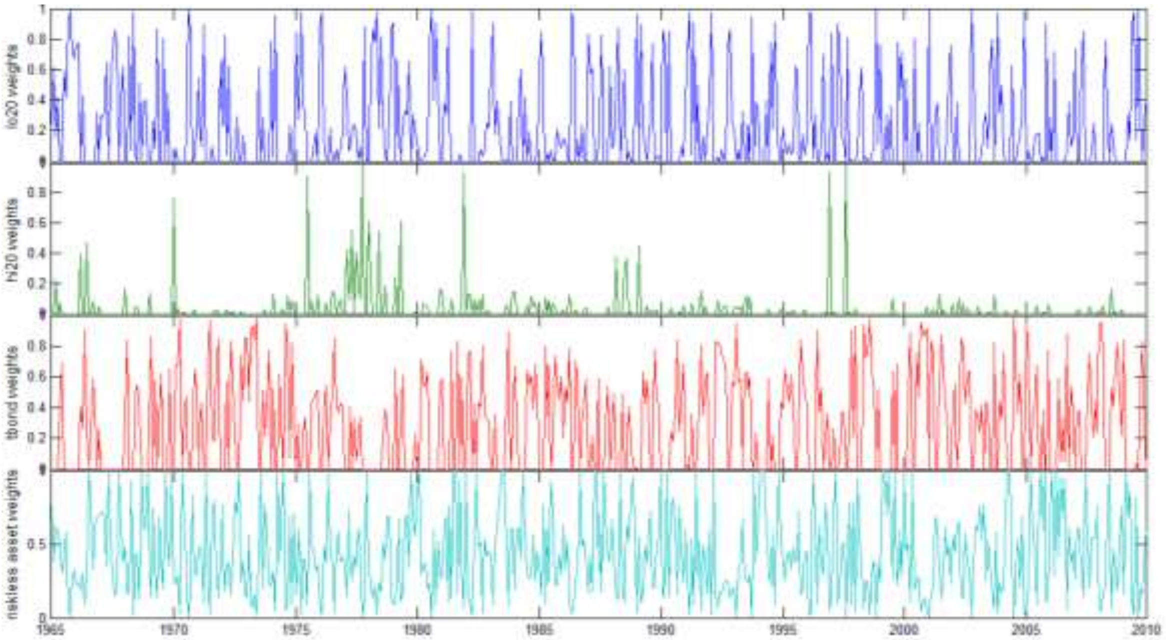


Figure 130 VAR(1) maximum Sharpe ratio portfolios weights (1 = 100%) with budget constraint (not permit to invest in the riskless asset)

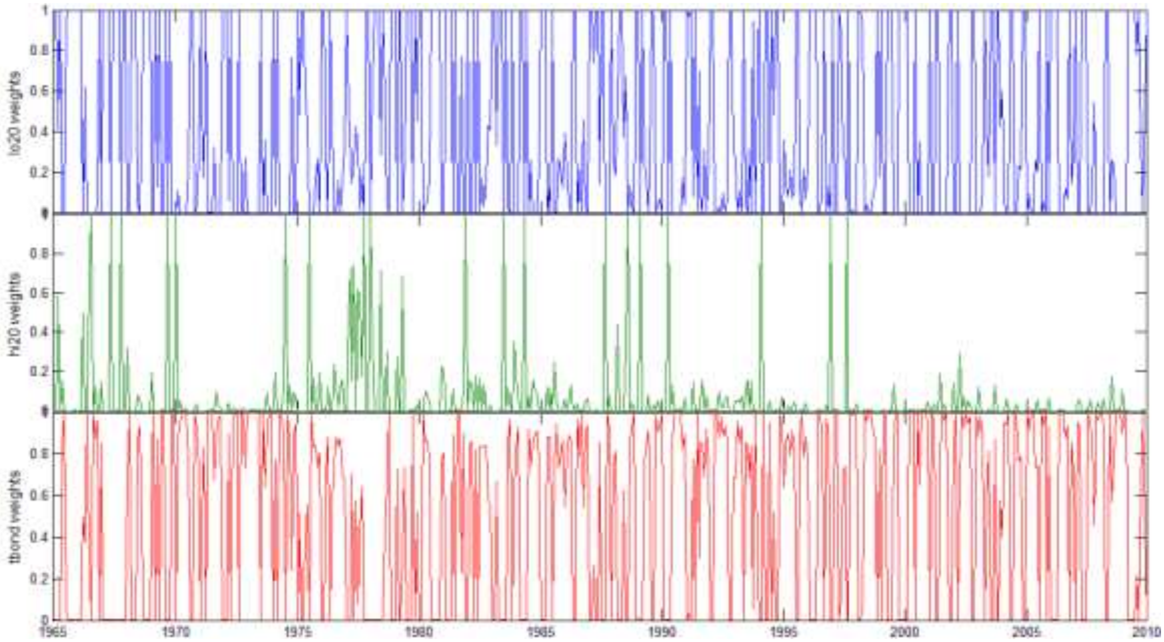


Figure 131 VAR(1) overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 1

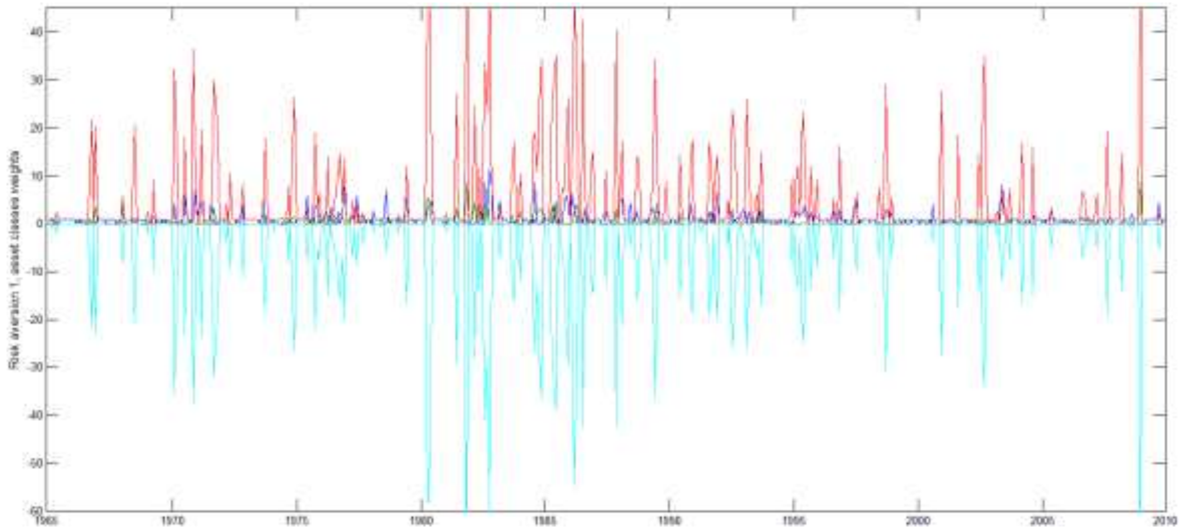


Figure 132 VAR(1) overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 3

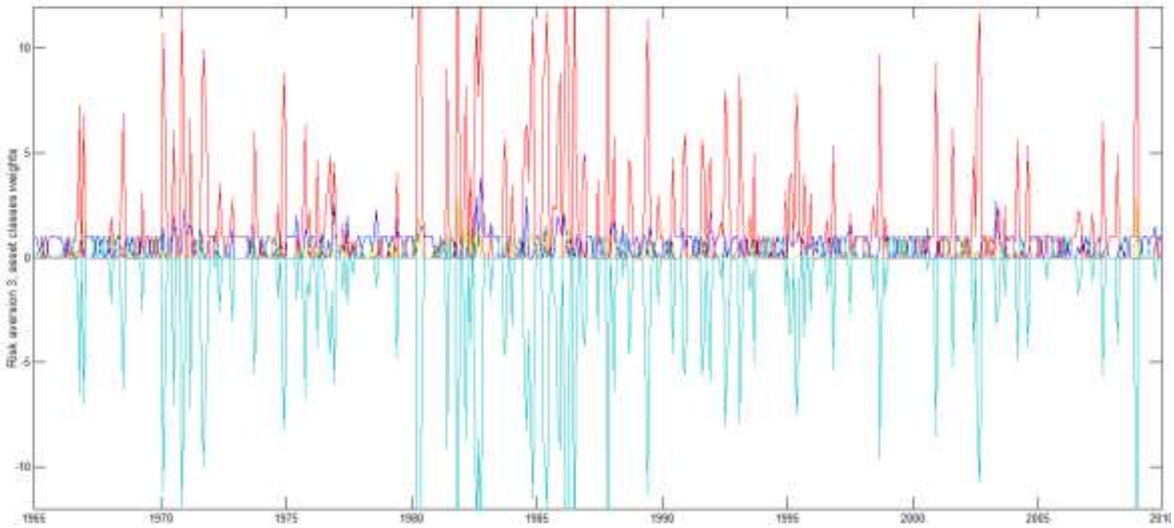


Figure 133 VAR(1) overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 5

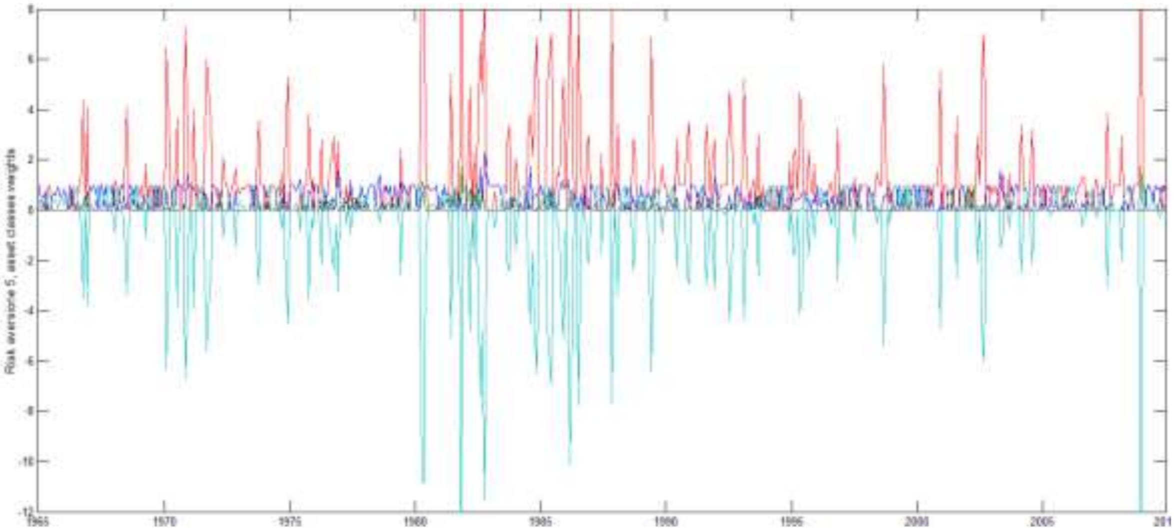


Figure 134 VAR(1) risky optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 1

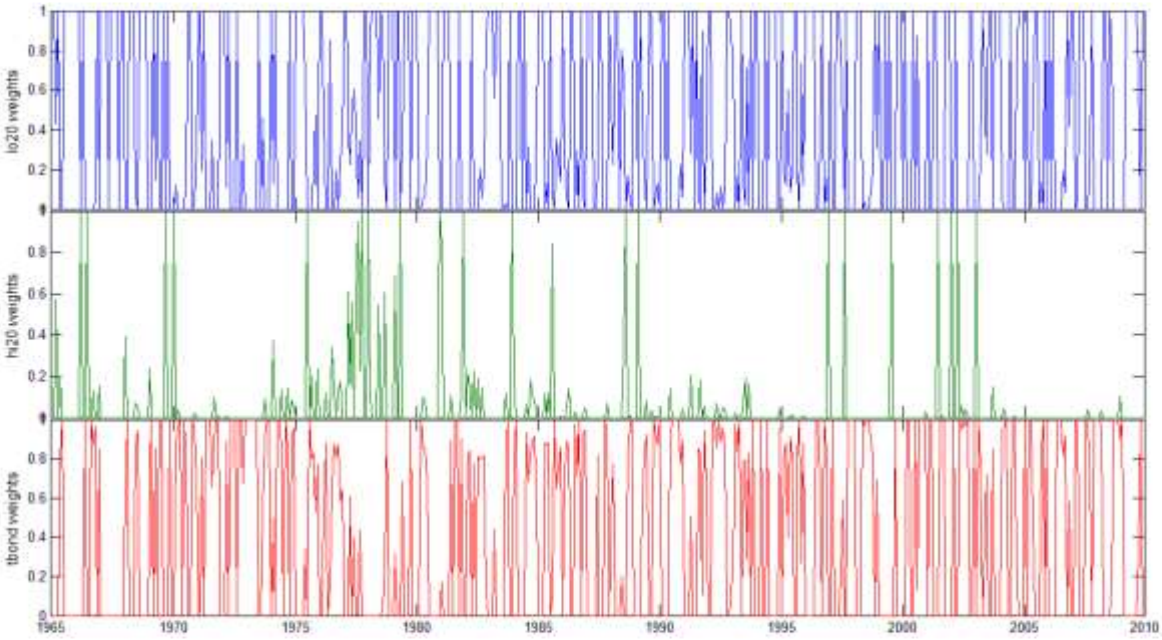


Figure 135 VAR(1) risky optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 3

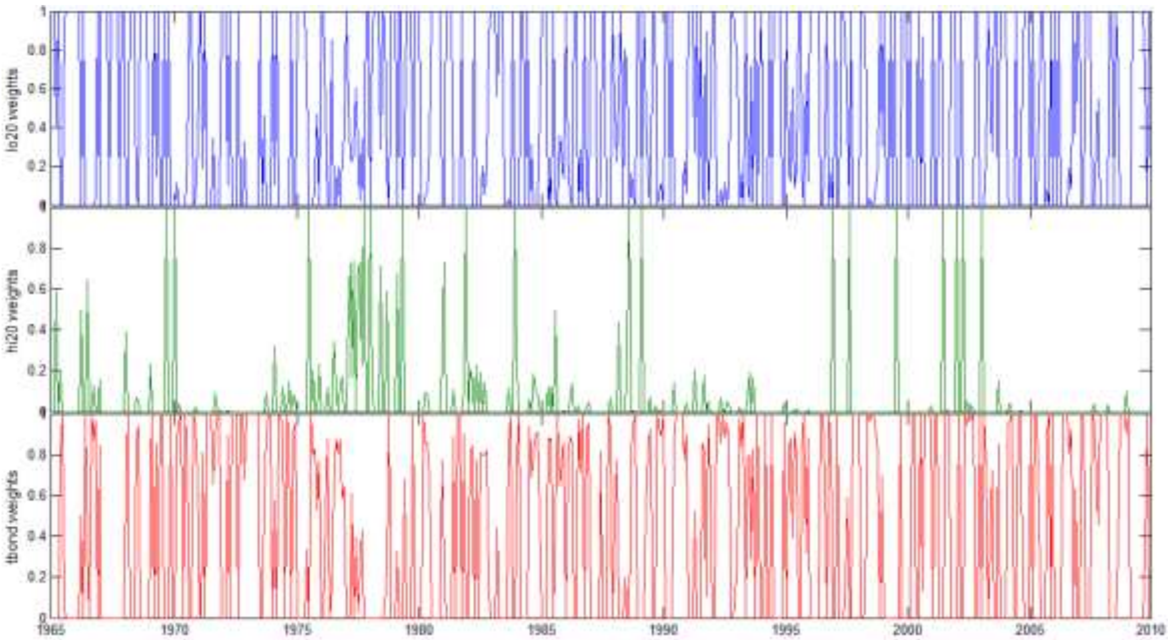


Figure 136 VAR(1) risky optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 5

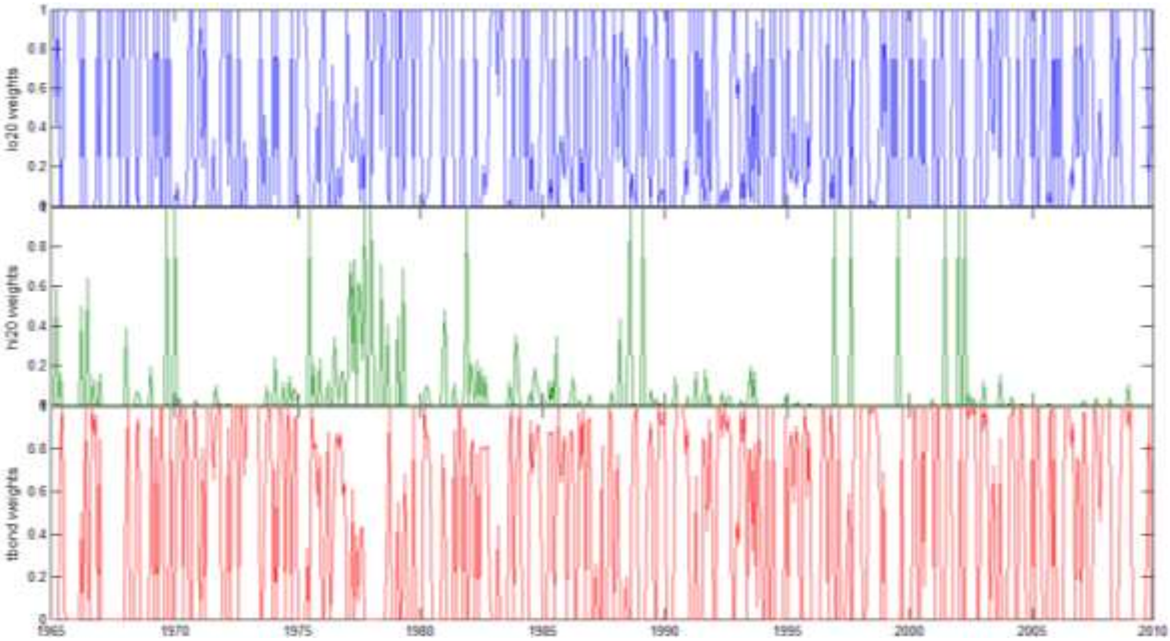


Table 26 VAR(1) in-sample average portfolio overall weights (1 = 100%)

avg. weights maximum Sharpe ratio portfolio with opened lower budget constraint							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.2384	(0.3162)	0.0361	(0.1220)	0.2876	(0.3118)	0.4380	(0.2953)
avg. weights maximum Sharpe ratio portfolio with budget constraint							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.4204	(0.4514)	0.0712	(0.2042)	0.5084	(0.4404)	0.0000	(0.0000)
avg. optimal weights with capital allocation and risk aversion coefficient = 1							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.9902	(1.5068)	0.3331	(0.9524)	4.4857	(9.7123)	-4.8090	(10.8207)
avg. optimal weights with capital allocation and risk aversion coefficient = 3							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.5321	(0.5839)	0.1268	(0.3394)	1.6242	(3.2006)	-1.2831	(3.4641)
avg. optimal weights with capital allocation and risk aversion coefficient = 5							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.3991	(0.4368)	0.0794	(0.2149)	1.0621	(1.9066)	-0.5406	(2.0231)

Table 27 VAR(1) in-sample average portfolio risky weights (1 = 100%)

avg. optimal risky weights with capital allocation and risk aversion

coefficient = 1			
lo20	*hi20*	*tbond*	*risk-free asset*
0.4470 (0.4645)	0.0662 (0.2108)	0.4164 (0.4492)	0.0000 (0.0000)

avg. optimal risky weights with capital allocation and risk aversion coefficient = 3			
lo20	*hi20*	*tbond*	*risk-free asset*
0.4415 (0.4621)	0.0646 (0.2016)	0.4235 (0.4470)	0.0000 (0.0000)

avg.optimal risky weights with capital allocation and risk aversion coefficient = 5			
lo20	*hi20*	*tbond*	*risk-free asset*
0.4343 (0.4592)	0.0595 (0.1866)	0.4358 (0.4454)	0.0000 (0.0000)

Table 28 VAR(1) in-sample portfolios expected moments

maximum Sharpe ratio portfolio with opened lower budget constraint		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0060 (0.0065)	0.0208 (0.0173)	0.2985 (0.2491)
maximum Sharpe ratio portfolio with budget constraint		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0097 (0.0097)	0.0372 (0.0202)	0.2859 (0.2701)
optimal weights with capital allocation and risk aversion coefficient = 1		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0677 (0.2105)	0.1396 (0.2049)	0.0808 (4.1512)
optimal weights with capital allocation and risk aversion coefficient = 3		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0276 (0.0699)	0.0623 (0.0639)	-0.2686 (12.4132)
optimal weights with capital allocation and risk aversion coefficient = 5		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0190 (0.0423)	0.0440 (0.0385)	-0.6195 (20.6788)

Figure 137 VAR(1) maximum Sharpe ratio portfolios (with and without budget constraint) expected returns

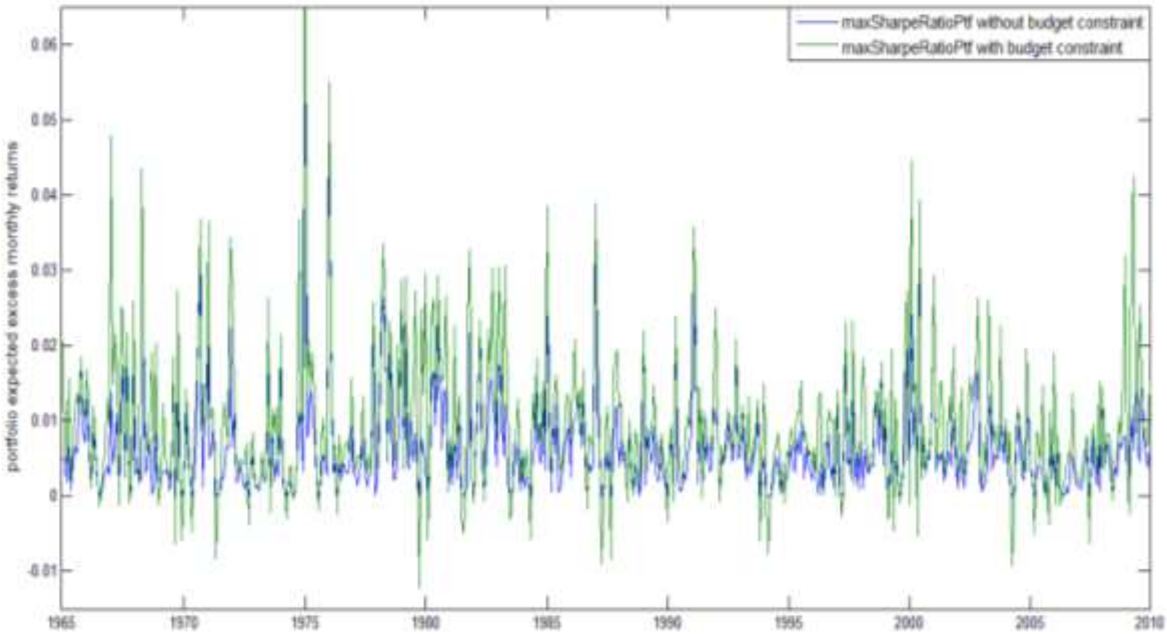


Figure 138 VAR(1) expected returns of the maximum expected utility portfolios (capped at 1 = 100%)

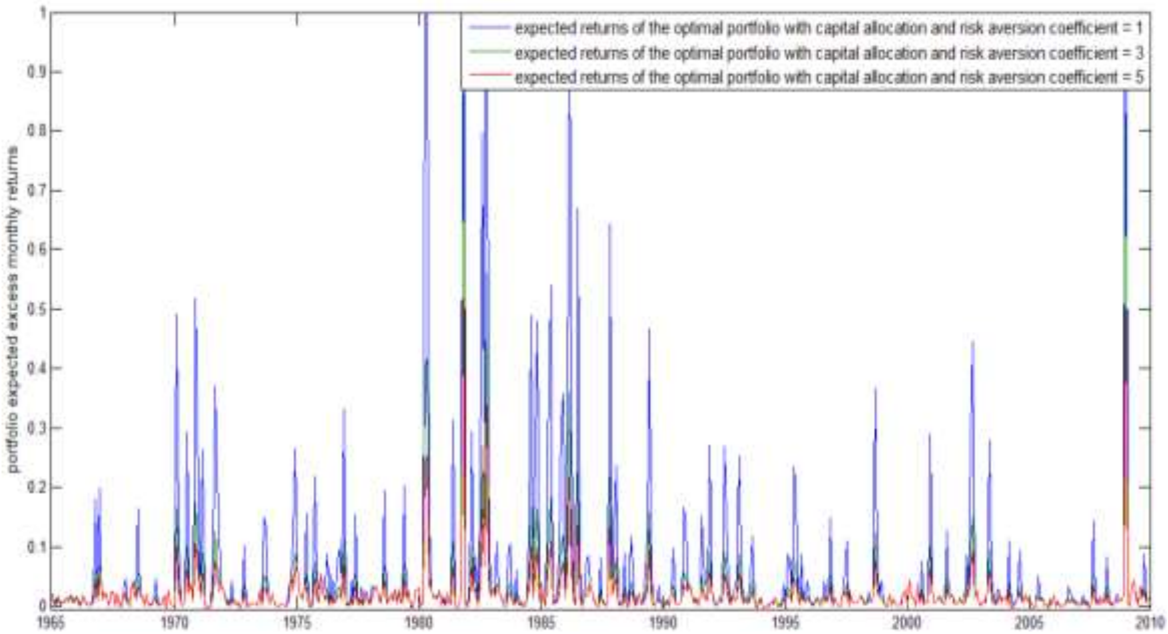


Table 29 VAR(1) in-sample realized portfolios statistics

maximum Sharpe ratio portfolio with opened lower budget constraint		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0153	0.0271	0.5649
maximum Sharpe ratio portfolio with budget constraint		

average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0236	0.0425	0.5546
optimal weights with capital allocation and risk aversion coefficient = 1		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.1824	0.6388	0.2855
optimal weights with capital allocation and risk aversion coefficient = 3		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0740	0.2135	0.3465
optimal weights with capital allocation and risk aversion coefficient = 5		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0510	0.1304	0.3909

Figure 139 VAR(1) in-sample maximum Sharpe ratio portfolio with opened lower budget constraint realized returns

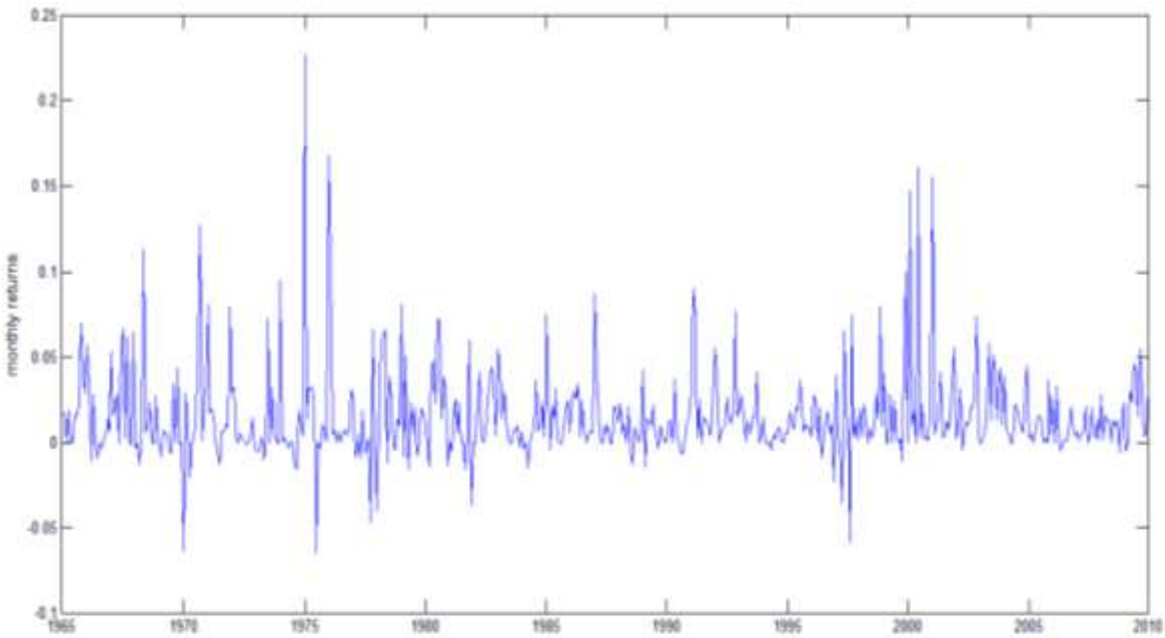


Figure 140 VAR(1) in-sample maximum Sharpe ratio portfolio with opened lower budget constraint cumulative realized returns

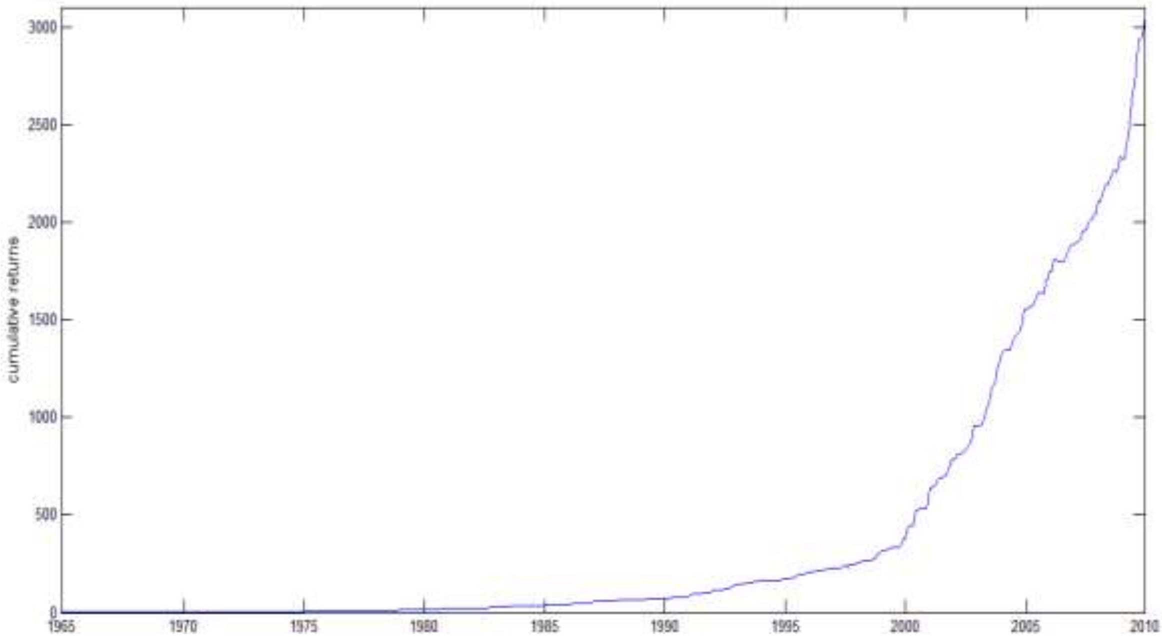


Figure 141 VAR(1) in-sample maximum Sharpe ratio portfolio with budget constraint realized returns

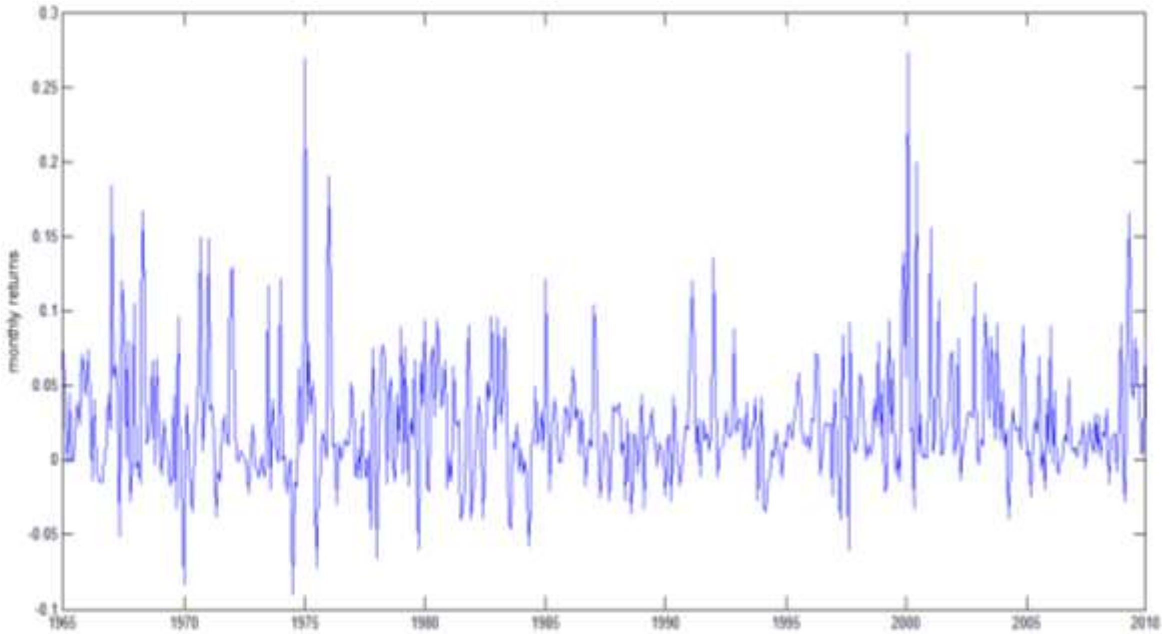


Figure 142 VAR(1) in-sample maximum Sharpe ratio portfolio with budget constraint realized cumulative returns

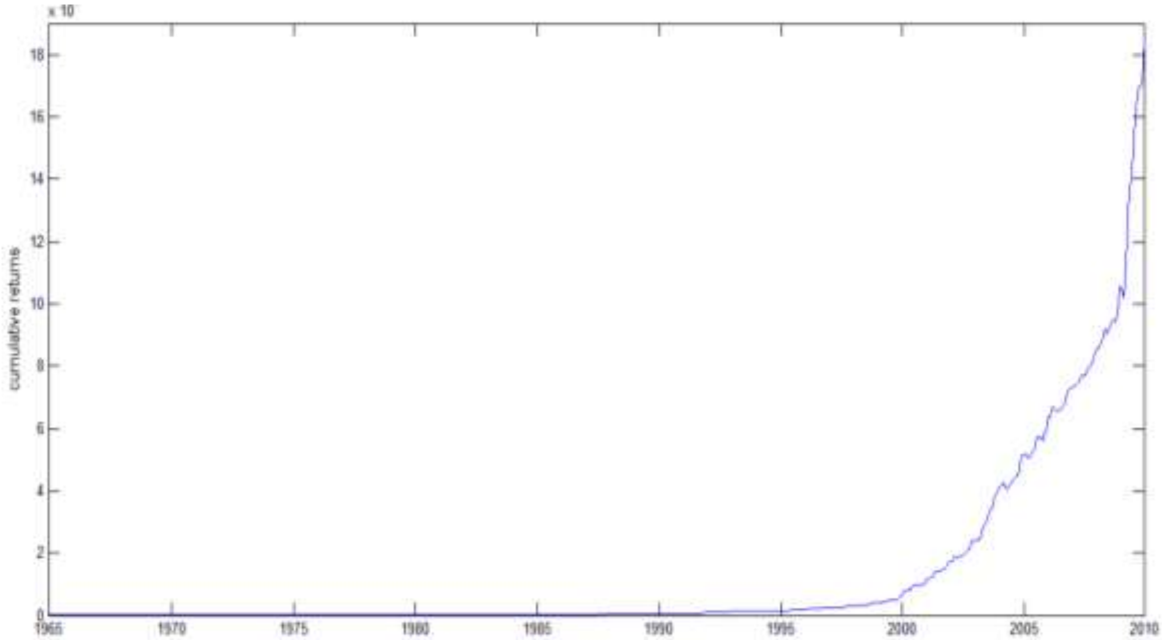


Figure 143 VAR(1) in-sample optimal portfolio with capital allocation and risk aversion coefficient = 1 realized returns

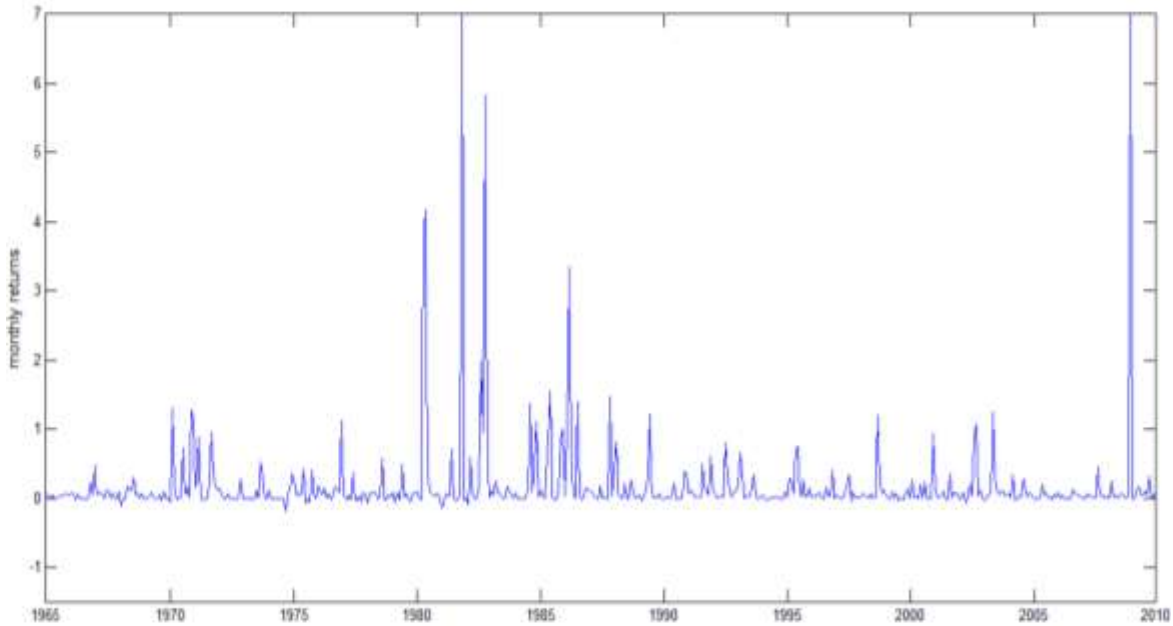


Figure 144 VAR(1) in-sample optimal portfolio with capital allocation and risk aversion coefficient = 1 cumulative realized returns

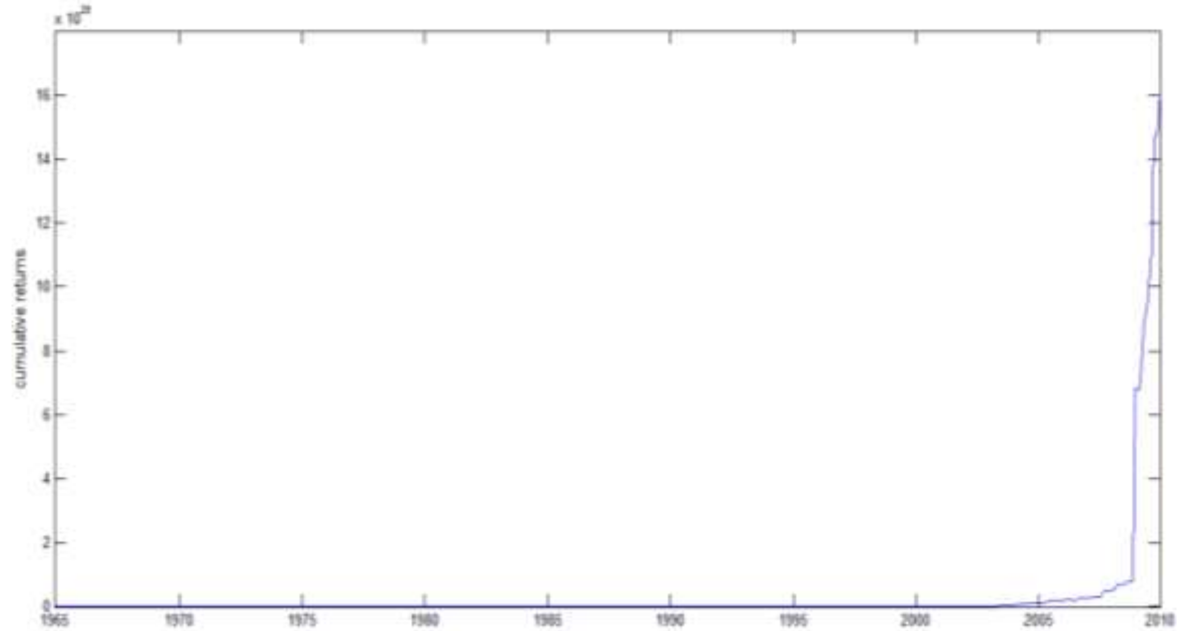


Figure 145   VAR(1) in-sample optimal portfolio with capital allocation and risk aversion coefficient = 3 realized returns

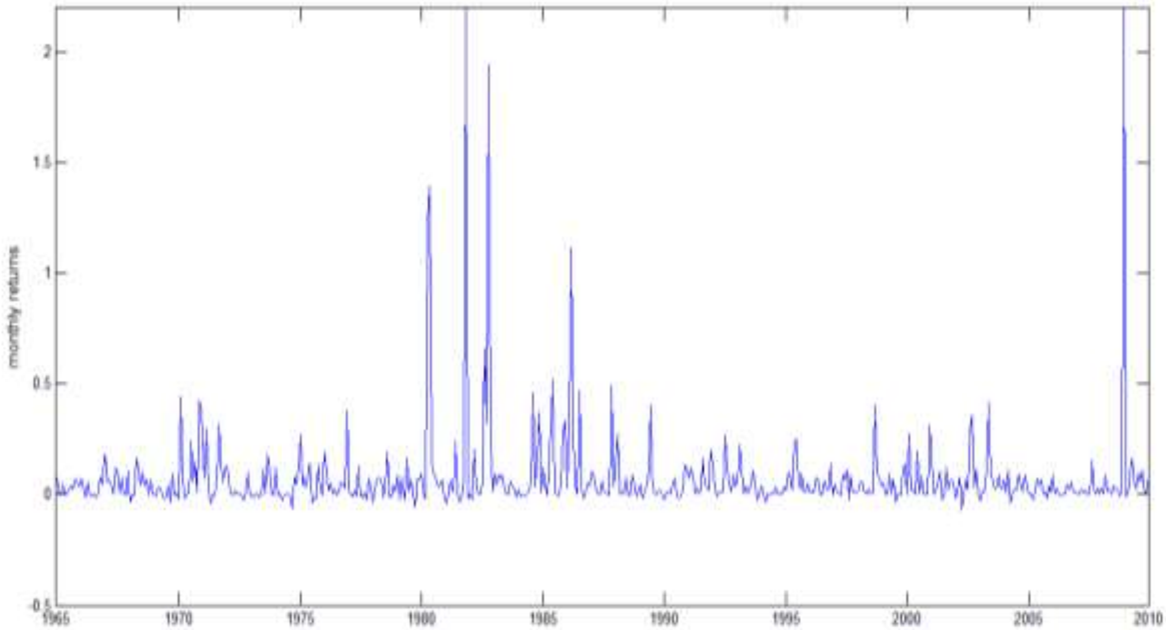


Figure 146   VAR(1) in-sample optimal portfolio with capital allocation and risk aversion coefficient = 3 cumulative realized returns

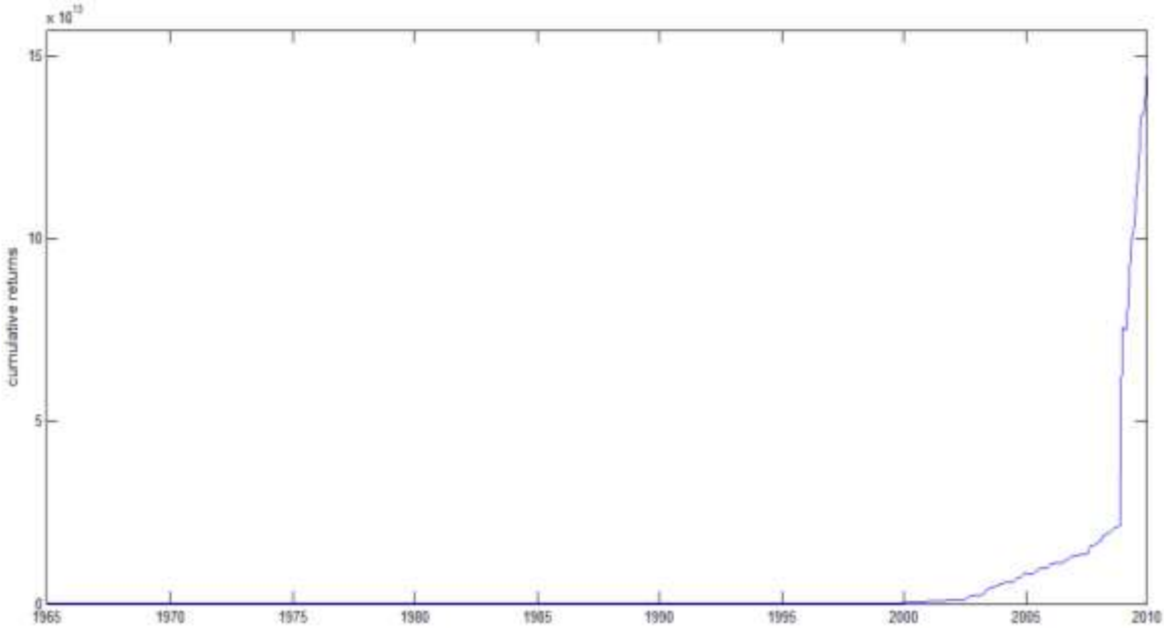


Figure 147 VAR(1) in-sample optimal portfolio with capital allocation and risk aversion coefficient = 5 realized returns

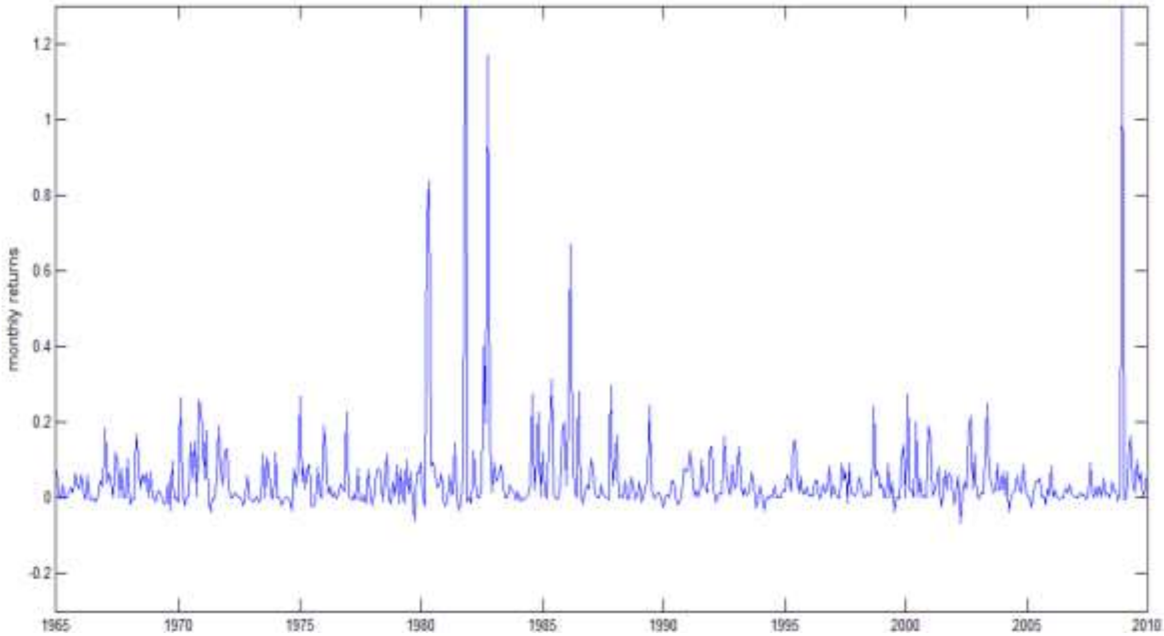
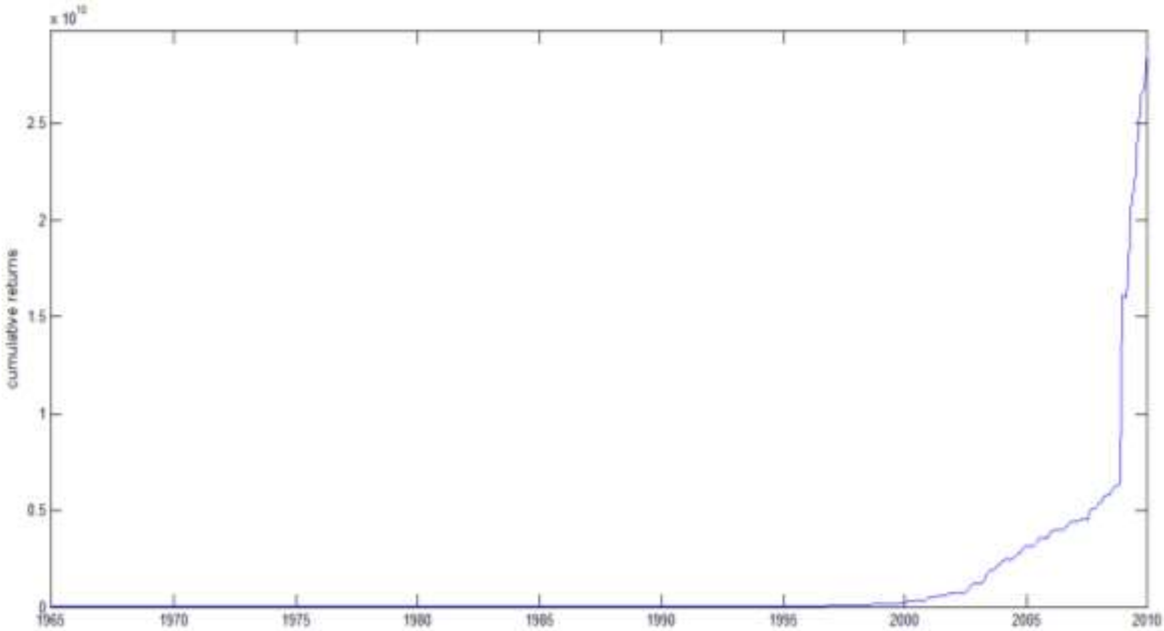


Figure 148 VAR(1) in-sample optimal portfolio with capital allocation and risk aversion coefficient = 5 cumulative realized returns



A comparison of the in-sample realized portfolio returns obtained using both the MSVAR(2,1) and the VAR(1) can be conducted from the observation of Table 25 and Table 29. Overall it emerges that none of the two models used to estimate the asset classes first two moments leads to a substantial portfolio overperformance over the other model. In fact, among the portfolios built using the MSVAR(2,1) estimates only the one that maximizes the expected utility function with risk aversion coefficient equal to 1 overperfoms the portfolios built using the more simple VAR(1) model; however if the comparison is conducted in terms of realized portfolio Sharpe ratio it can be seen that the portfolios built using the two states model overperform those built using the single state model two out of three times.

4.3 Out-of-sample asset allocation exercise

In this paragraph I am going to illustrate the out-of-sample portfolio construction processes that have been adopted and both the realized and the expected asset allocation and portfolio statistics, using firstly the MSVAR(2,1) and secondly the VAR(1) to model the first two moments of the asset classes returns distributions. The out-of-sample period starts at January 2010, lasts 5 years and end at December 2014. For the portfolios built using the asset classes first two expected moments estimated from a MSVAR(2,1) model and from a VAR(1) model in the first paragraph the portfolios weights results are shown as well as some expected

portfolio statistics while in the second subparagraph the portfolio realized returns are exhibited and commented.

4.3.1 *Out-of-sample asset allocation exercise based on a MSVAR(2,1) model*

The out-of-sample exercise required me to estimate the asset classes first two expected moments using the MSVAR(2,1) already estimated. Although the two state conditional VAR matrices regressive and autoregressive coefficients have been estimated using the in-sample returns, from 1965 to 2010, the time *t* filtered probabilities in this out-of-sample exercise are no more a by-product of the likelihood function maximization, indeed they must be calculated with the Hamilton's filter updating prior beliefs regard them, formulated at time *t-1,* with the arrival of new information at time *t* according to the following formula:

(55) $$\hat{\boldsymbol{\xi}}_{\mathbf{t|t}} = \frac{\mathbf{P}\hat{\boldsymbol{\xi}}_{\mathbf{t-1|t-1}} \odot \boldsymbol{\eta_t}}{f(\boldsymbol{R_t}|\Im_{\mathrm{t-1}};\boldsymbol{\theta})}$$

where $\boldsymbol{\xi}_{\mathbf{t|t}}$ is a *(1 x 2)* vector and **P** is *(2 x 2)* matrix that governs the transition dynamic (its coefficients have been estimated using the in-sample returns), $\boldsymbol{\eta_t}$ is a *(2 x 1)* vector whose *j*-th element is the density in regime *j* of the observation $\boldsymbol{R_t}$:

(56) $$\boldsymbol{\eta_{jt}} = f(\boldsymbol{R_t}|\mathrm{S_t} = j, \Im_{\mathrm{t-1}};\boldsymbol{\theta})$$

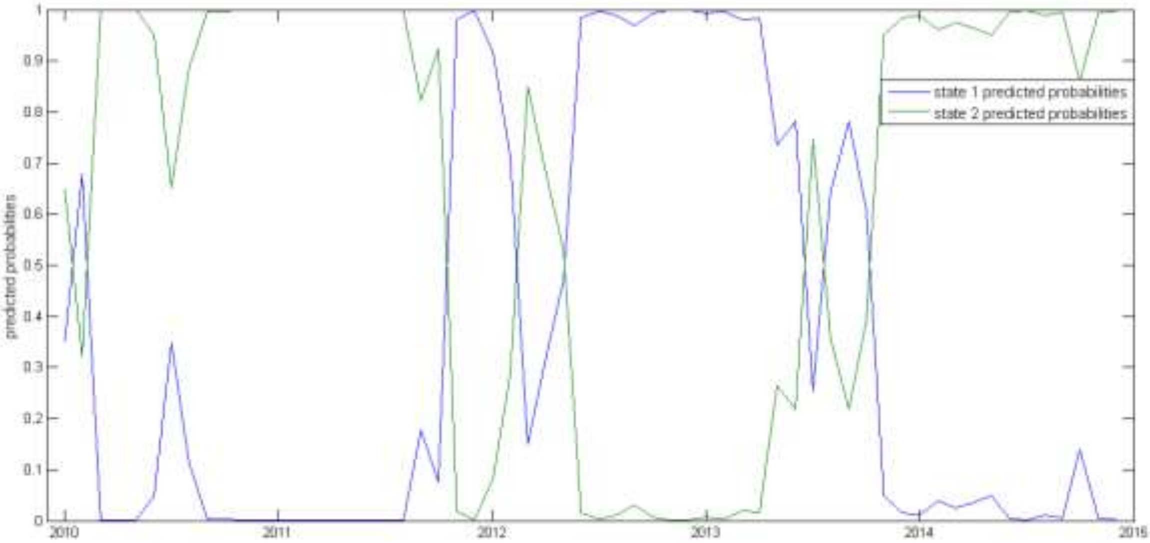
while the following formula is the unconditional density of the observation $\boldsymbol{R_t}$**:**

(57)[22] $$f(\boldsymbol{R_t}|\Im_{\mathrm{t-1}};\boldsymbol{\theta}) = \mathbf{1}'\left(\mathbf{P}\hat{\boldsymbol{\xi}}_{\mathbf{t-1|t-1}} \odot \boldsymbol{\eta_t}\right)$$

In the out-of-sample exercise that I perform, the last available estimated and not updated filtered probability is $\boldsymbol{\xi}_{\mathbf{540}}$ relative to December 2009. As a consequence any next probability in the out-of-sample exercise has been estimated with the Hamilton's filter according to formula (55).

Figure 149 out-of-sample predicted probabilities

[22] James D. Hamilton "Regime-Switching Models. Palgrave Dictionary of Economics" (2005) and Birger Nilsson and Andreas Graflund "Dynamic Portfolio Selection: The Relevance of Switching Regimes and Investment Horizon" (2001)



Given that the investor is not interested in investing in the dividend yield but only in the small stocks, large stocks and bonds, the predicted dynamic conditional mean for each t+T from January 2010 to December 2014 is equal to the first three elements of the following vector:

(58) $\mathrm{E}(\boldsymbol{y}_{\mathbf{t+T}}|\boldsymbol{y}_{\mathbf{t+T-1}};\boldsymbol{\theta};\Im_{\mathrm{t+T-1}}) = \hat{\boldsymbol{\mu}}_{\mathrm{t+T|t+T-1}} = \xi_{1,\mathrm{t+T|t+T-1}}\mathrm{E}(\boldsymbol{y}_{\mathbf{t+T}}|\boldsymbol{y}_{\mathbf{t+T-1}};\mathrm{S_t}=1;\boldsymbol{\theta};\Im_{\mathrm{t+T-1}}) + \xi_{2,\mathrm{t+T|t+T-1}}\mathrm{E}(\boldsymbol{y}_{\mathbf{t+T}}|\boldsymbol{y}_{\mathbf{t+T-1}};\mathrm{S_t}=2;\boldsymbol{\theta};\Im_{\mathrm{t+T-1}})$

(59) $\mathrm{E}(\boldsymbol{y}_{\mathbf{t+T}}|\boldsymbol{y}_{\mathbf{t+T-1}};\mathrm{S_t}=j;\boldsymbol{\theta};\Im_{\mathrm{t+T-1}}) = \hat{\boldsymbol{\mu}}_{\mathrm{t+T|t+T-1};\mathbf{S_t}=\boldsymbol{j}} = \boldsymbol{\mu}_{\mathbf{S_t}=\boldsymbol{j}} + \boldsymbol{A}_{\mathbf{1},\mathbf{S_t}=\boldsymbol{j}}\boldsymbol{y}_{\mathbf{t+T-1}}$

As it can be seen, similarly to what have been done for the expected in-sample vector of dynamic conditional mean, the one step ahead prediction *(t+1)* uses the current returns *(t)*; it follows that the time *t+T* prediction, which for example might be the prediction for December 2011 (t=540 that corresponds to December 2009 and T=12) uses the November 2011 (t=540 that corresponds to December 2009 and T=11) realized returns. Here $\boldsymbol{\mu}_{\mathbf{S_t}=\boldsymbol{j}}$ is the vector of intercept terms of $\boldsymbol{y}_{\mathbf{t+T}}$ in state $\mathrm{S_t}=j$, $\boldsymbol{A}_{\boldsymbol{j},\mathbf{S_t}=\boldsymbol{j}}$ is the matrix of autoregressive and regression coefficients associated with lag 1 in state $\mathrm{S_t}=j$; their values have been estimated using the in-sample returns.

The predicted dynamic variance-covariance matrix for each *t+T* from January 2010 to December 2014 is equal to the *(3 x 3)* left-upper part of the following matrix:

(60) $Var(\boldsymbol{y}_{\mathbf{t+T}}|\boldsymbol{y}_{\mathbf{t+T-1}};\boldsymbol{\theta};\Im_{\mathrm{t+T-1}}) = \xi_{1,\mathrm{t+T|t+T-1}}\boldsymbol{\Sigma}_{\mathbf{1}} + \xi_{2,\mathrm{t+T|t+T-1}}\boldsymbol{\Sigma}_{\mathbf{2}} +$
$\xi_{1,\mathrm{t+T|t+T-1}}\xi_{2,\mathrm{t+T|t+T-1}}(\widehat{\boldsymbol{\mu}}_{\mathrm{t+T|t+T-1};\mathbf{S_t=1}} - \widehat{\boldsymbol{\mu}}_{\mathrm{t+T|t+T-1};\mathbf{S_t=2}})(\widehat{\boldsymbol{\mu}}_{\mathrm{t+T|t+T-1};\mathbf{S_t=1}} -$
$\widehat{\boldsymbol{\mu}}_{\mathrm{t+T|t+T-1};\mathbf{S_t=2}})^{\boldsymbol{T}}$

where $\widehat{\boldsymbol{\mu}}_{\mathrm{t+T|t+T-1};\mathbf{S_t=1}}$ and $\widehat{\boldsymbol{\mu}}_{\mathrm{t+T|t+T-1};\mathbf{S_t=2}}$ are the predicted dynamic conditional means calculated above.

Figure 150 out-of-sample lo20 and its predicted dynamic conditional mean

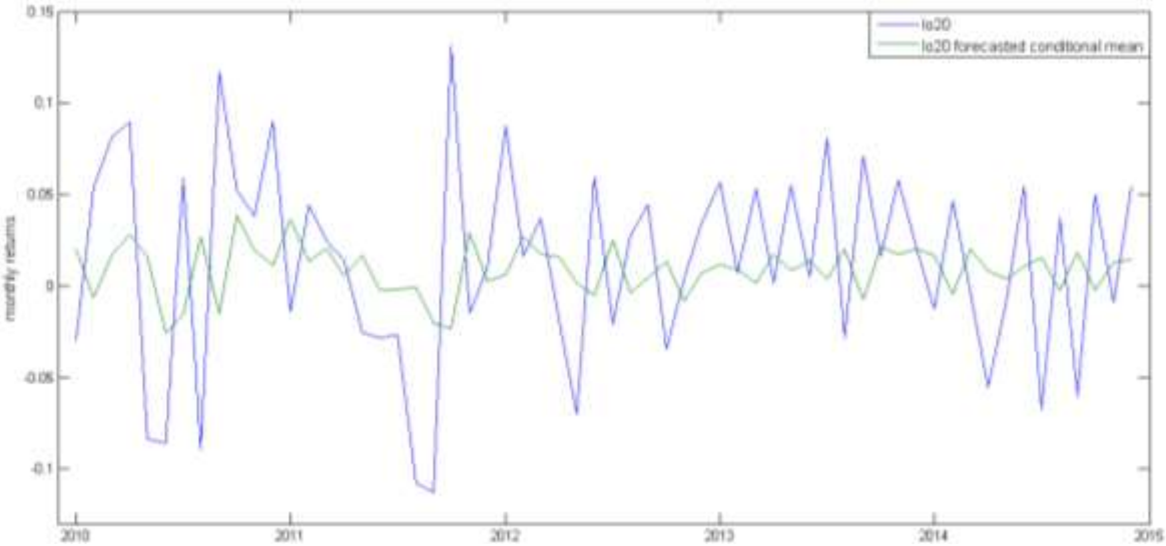


Figure 151 out-of-sample hi20 and its predicted dynamic conditional mean

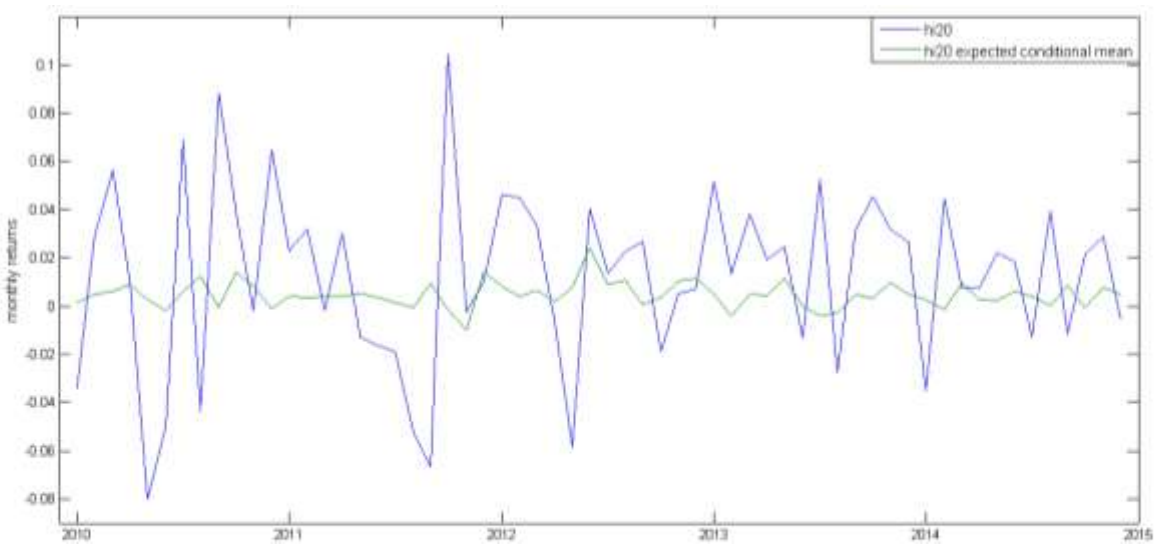
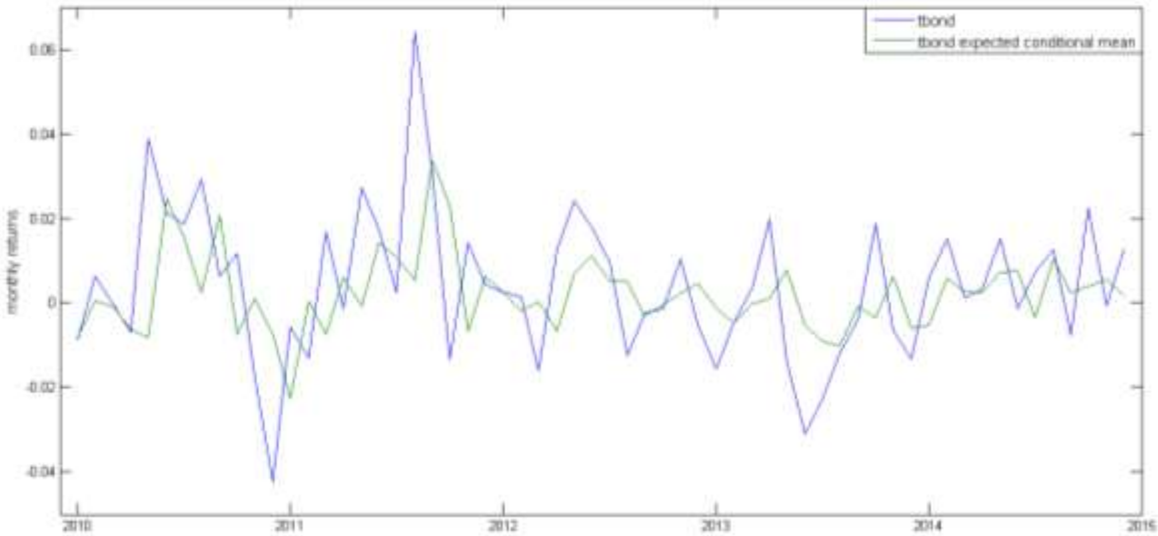


Figure 152 out-of-sample tbond and its predicted dynamic conditional mean



I then used the predicted dynamic variance-covariance matrix and the predicted dynamic conditional mean vector to build the same five in-sample recursive optimal portfolios. The results are here illustrated.

Figure 153 out-of-sample maximum Sharpe ratio portfolios weights (1 = 100%) with opened lower budget constraint (permit to invest in the riskless asset)

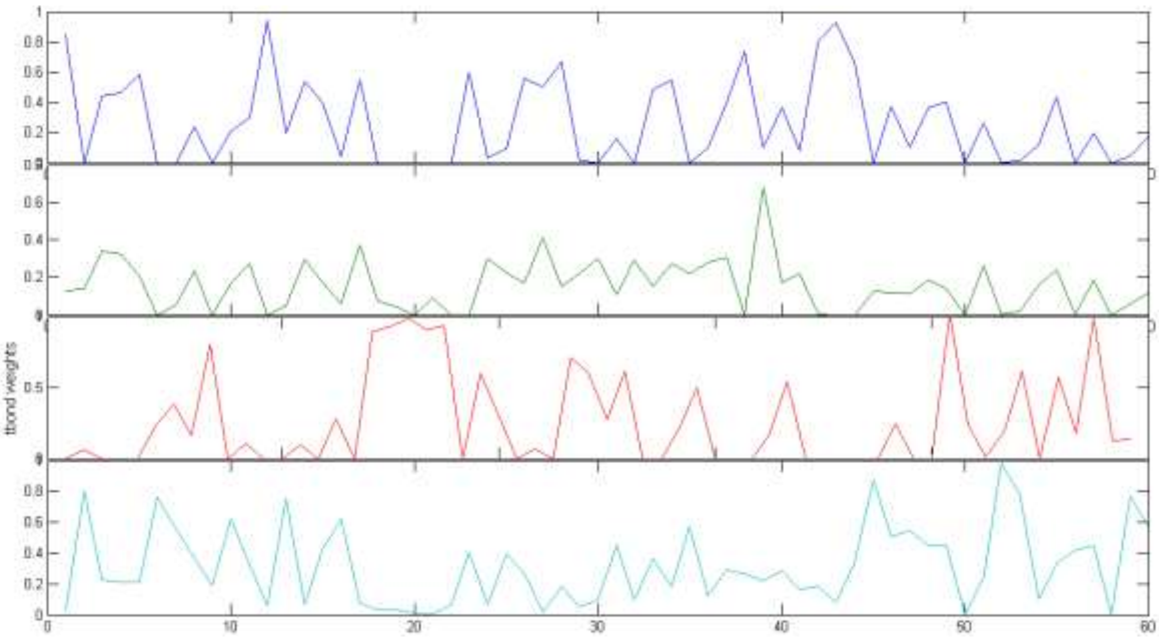


Figure 154 out-of-sample maximum Sharpe ratio portfolios weights (1 = 100%) with budget constraint (not permit to invest in the riskless asset)

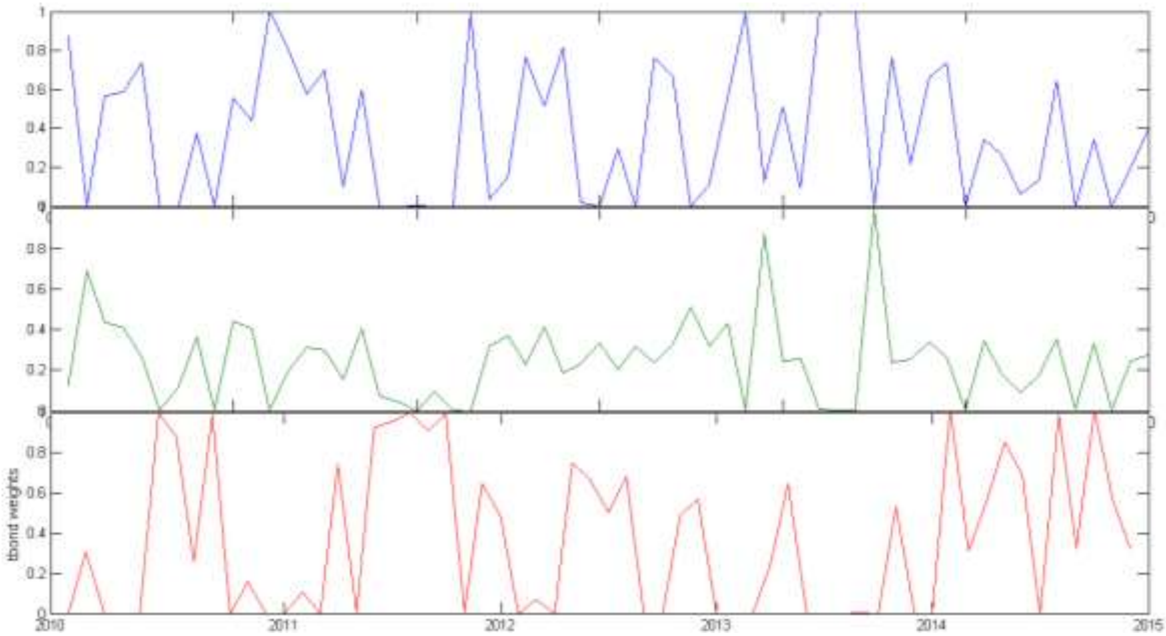


Figure 155 out-of-sample overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 1

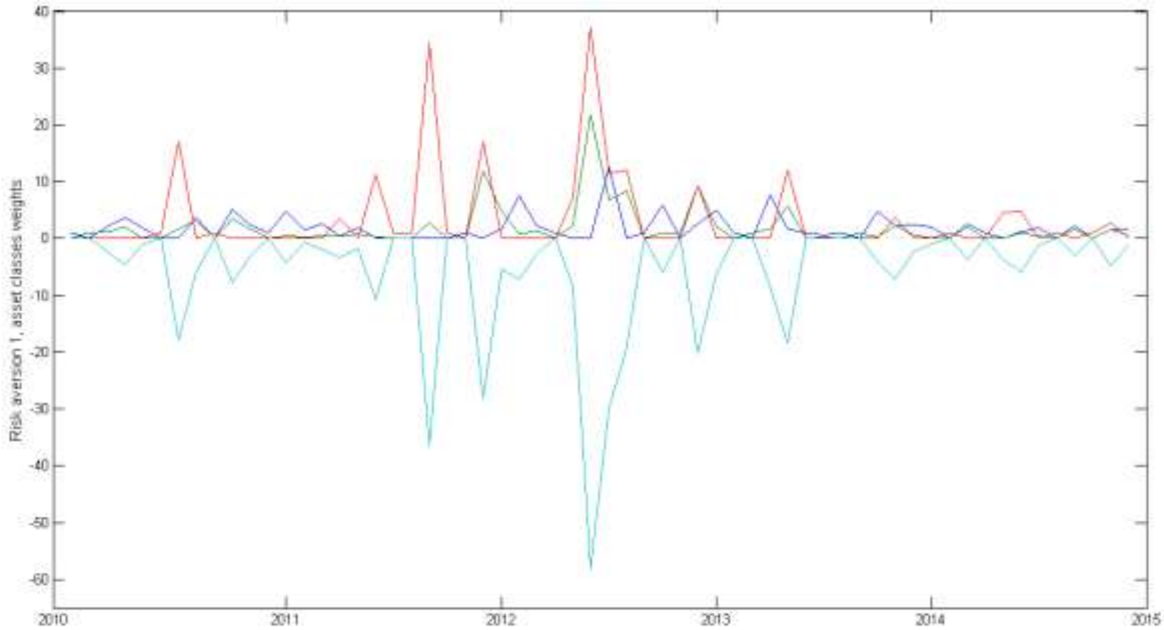


Figure 156 out-of-sample overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 3

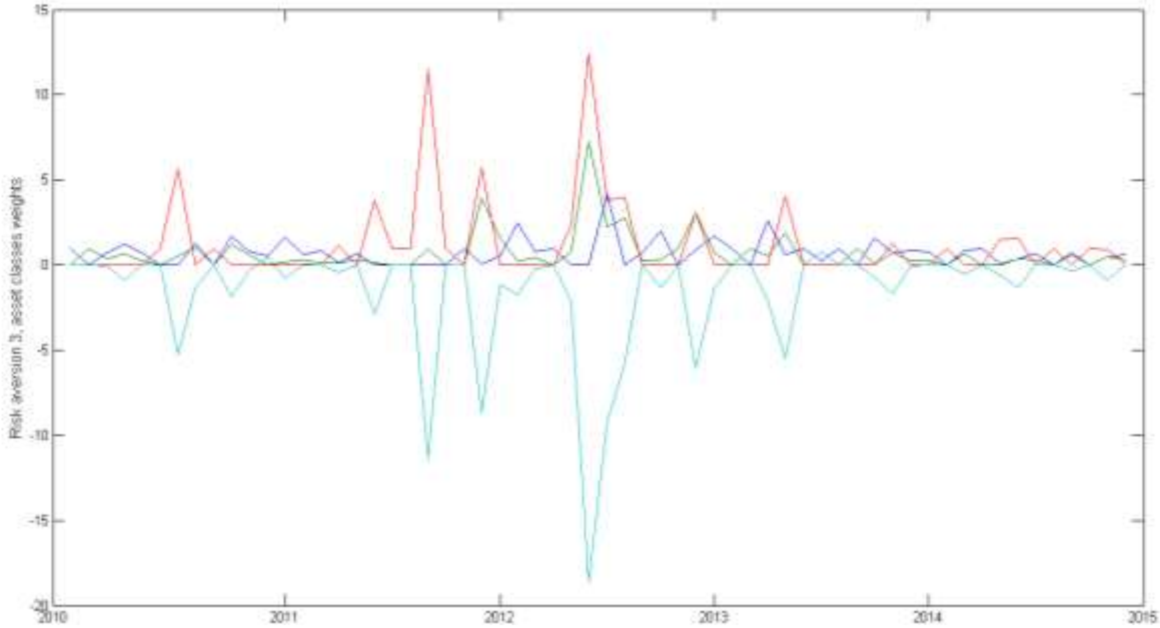


Figure 157 out-of-sample overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 5

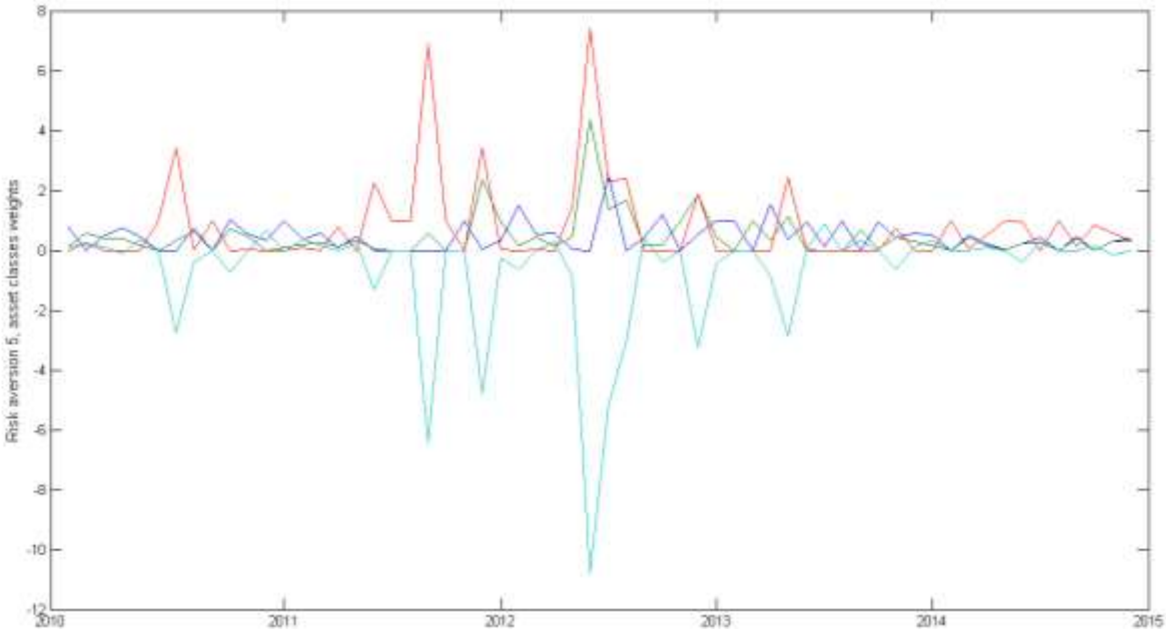


Table 30 out-of-sample average portfolio overall weights (1 = 100%)

avg. weights maximum Sharpe ratio portfolio with opened lower budget			
lo20	*hi20*	*tbond*	*risk-free asset*
0.2677 (0.2783)	0.1535 (0.1346)	0.2619 (0.3331)	0.3169 (0.2565)

avg. weights maximum Sharpe ratio portfolio with budget constraint

lo20		hi20		tbond		risk-free asset	
0.3852	(0.3536)	0.2459	(0.2062)	0.3689	(0.3858)	0.0000	(0.0000)

avg. optimal weights with capital allocation and risk aversion coefficient = 1

lo20		hi20		tbond		risk-free asset	
1.8252	(2.3385)	1.8493	(3.5482)	3.2800	(7.4712)	-5.9545	(10.4384)

avg. optimal weights with capital allocation and risk aversion coefficient = 3

lo20		hi20		tbond		risk-free asset	
0.7042	(0.7785)	0.6822	(1.1746)	1.1844	(2.4722)	-1.5707	(3.3469)

avg. optimal weights with capital allocation and risk aversion coefficient = 5

lo20		hi20		tbond		risk-free asset	
0.4541	(0.4891)	0.4496	(0.7038)	0.7851	(1.4767)	-0.6888	(1.9459)

Table 31 out-of-sample average portfolio risky weights (1 = 100%)

avg. optimal risky weights with capital allocation and risk aversion coefficient=1

lo20		hi20		tbond		risk-free asset	
0.4721	(0.4136)	0.2242	(0.2708)	0.3036	(0.4114)	0.0000	(0.0000)

avg. optimal risky weights with capital allocation and risk aversion coefficient = 3

lo20		hi20		tbond		risk-free asset	
0.4520	(0.3960)	0.2422	(0.2665)	0.3058	(0.4099)	0.0000	(0.0000)

avg. optimal risky weights with capital allocation and risk aversion coefficient = 5

lo20		hi20		tbond		risk-free asset	
0.4169	(0.3738)	0.2550	(0.2509)	0.3281	(0.4023)	0.0000	(0.0000)

Figure 158 out-of-sample maximum Sharpe ratio portfolios (with and without budget constraint) expected returns

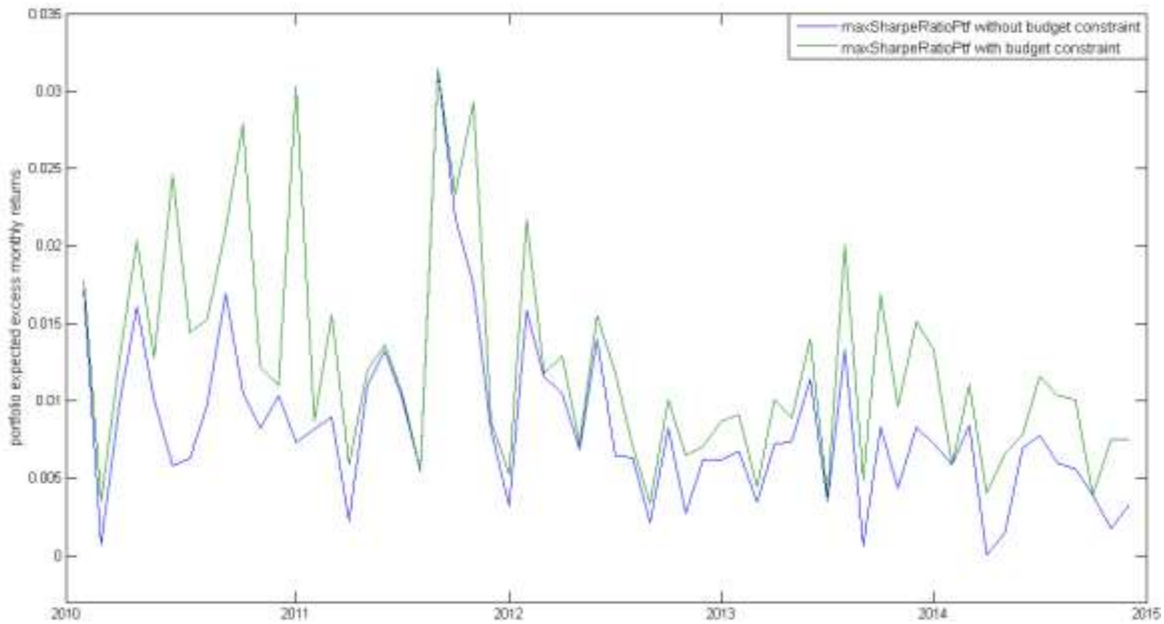


Figure 159 out-of-sample expected returns of the maximum expected utility portfolios (capped at 1 = 100%)

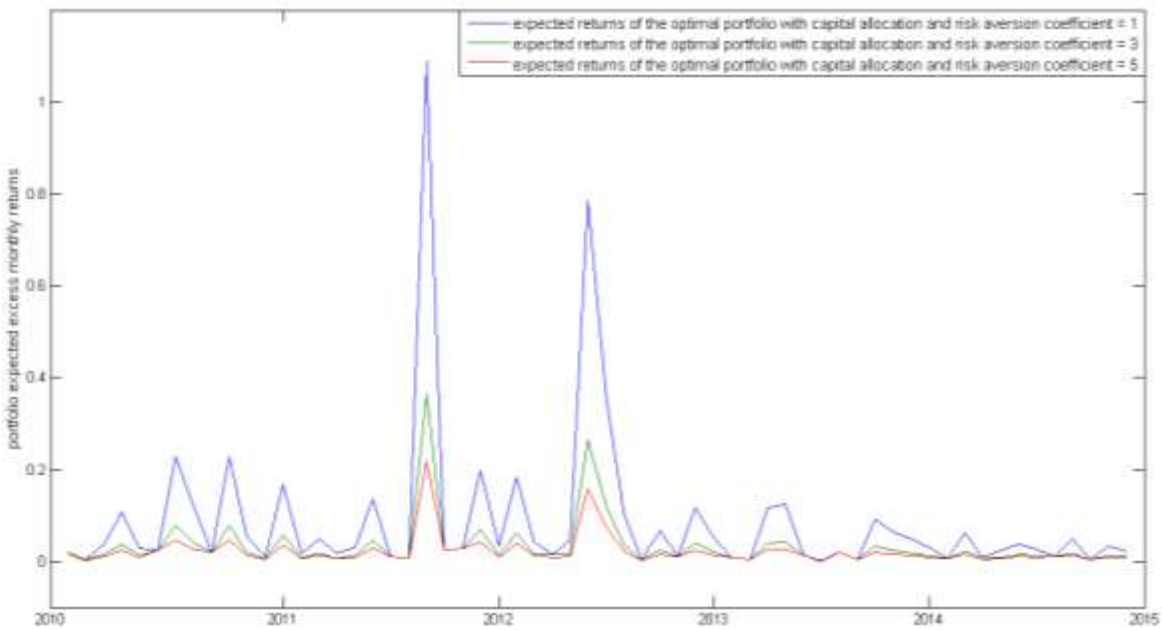


Table 32 out-of-sample portfolios expected moments

maximum Sharpe ratio portfolio with opened lower budget constraint		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0083 (0.0055)	0.0266 (0.0164)	0.3559 (0.2220)
maximum Sharpe ratio portfolio with budget constraint		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*

0.0122 (0.0070)	0.0386 (0.0169)	0.3559 (0.2220)
optimal weights with capital allocation and risk aversion coefficient = 1		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0884 (0.1756)	0.2100 (0.1994)	0.3075 (0.2013)
optimal weights with capital allocation and risk aversion coefficient = 3		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0335 (0.0578)	0.0806 (0.0595)	0.3208 (0.1982)
optimal weights with capital allocation and risk aversion coefficient = 5		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0222 (0.0346)	0.0527 (0.0337)	0.3305 (0.1973)

In this section of the subparagraph the out-of-sample portfolio realized results are shown.

Table 33 out-of-sample realized portfolios statistics

maximum Sharpe ratio portfolio with opened lower budget constraint		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0084	0.0246	0.3406
maximum Sharpe ratio portfolio with budget constraint		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0115	0.0305	0.3788
optimal weights with capital allocation and risk aversion coefficient = 1		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0729	0.2437	0.2990
optimal weights with capital allocation and risk aversion coefficient = 3		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0266	0.0832	0.3203
optimal weights with capital allocation and risk aversion coefficient = 5		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0181	0.0509	0.3555

Figure 160 out-of-sample maximum Sharpe ratio portfolio with opened lower budget constraint realized returns

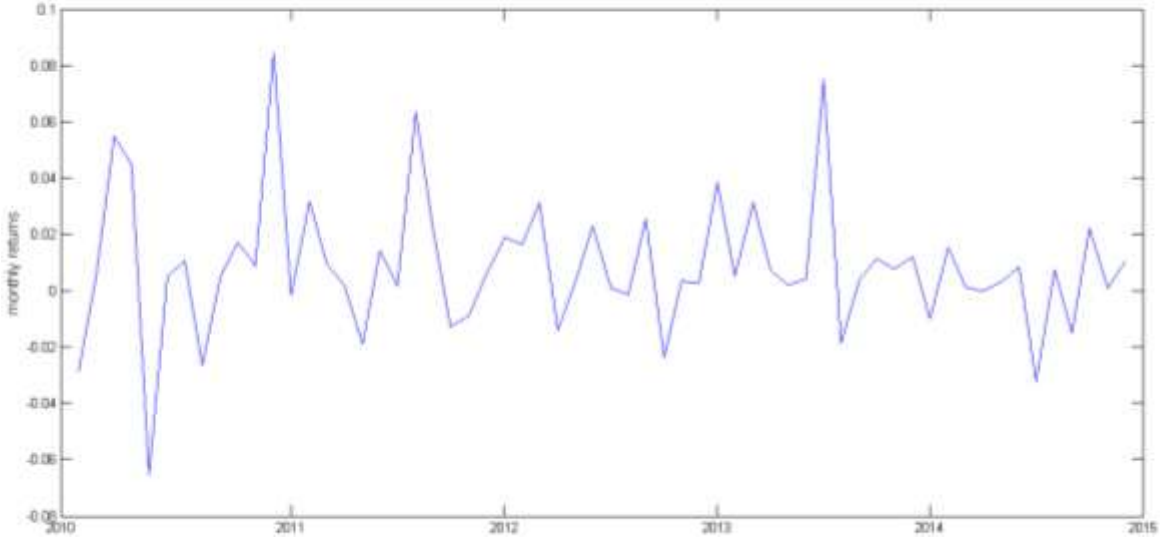


Figure 161 out-of -sample maximum Sharpe ratio portfolio with opened lower budget constraint cumulative realized returns

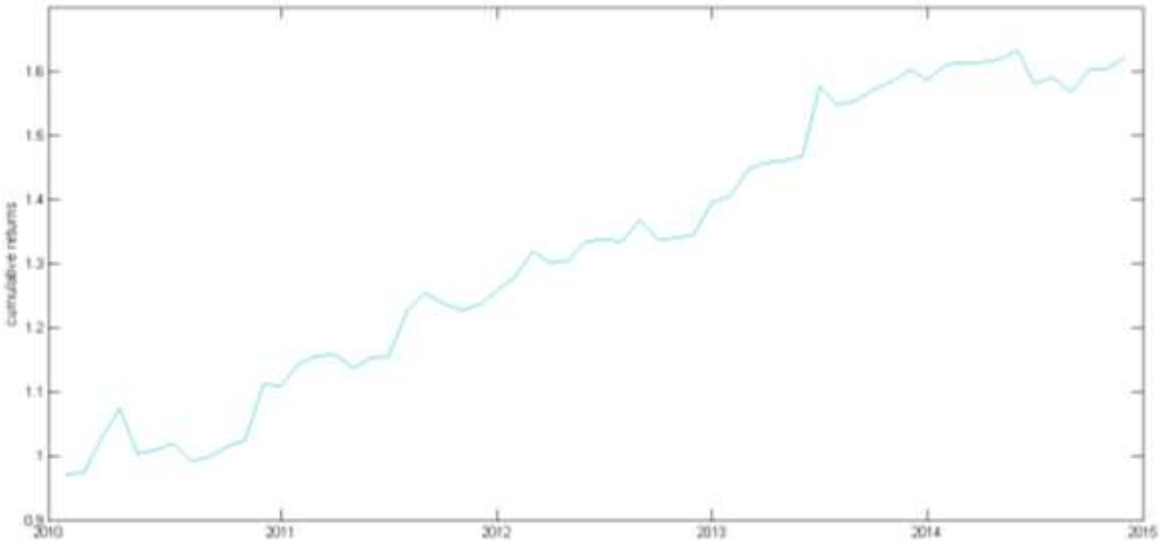


Figure 162 out-of -sample maximum Sharpe ratio portfolio with budget constraint realized returns

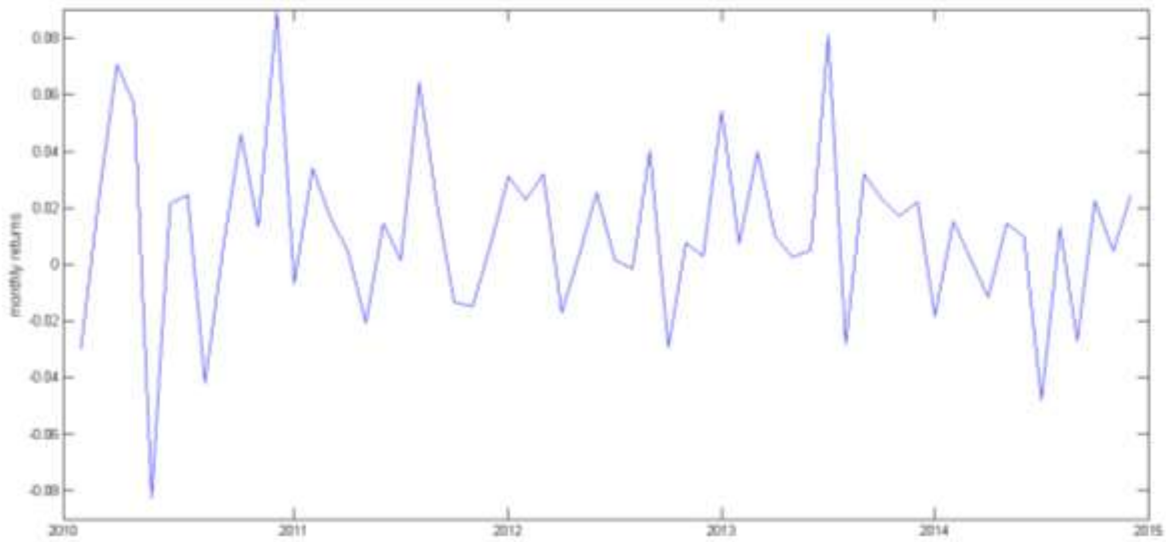


Figure 163   out-of -sample maximum Sharpe ratio portfolio with budget constraint realized cumulative returns

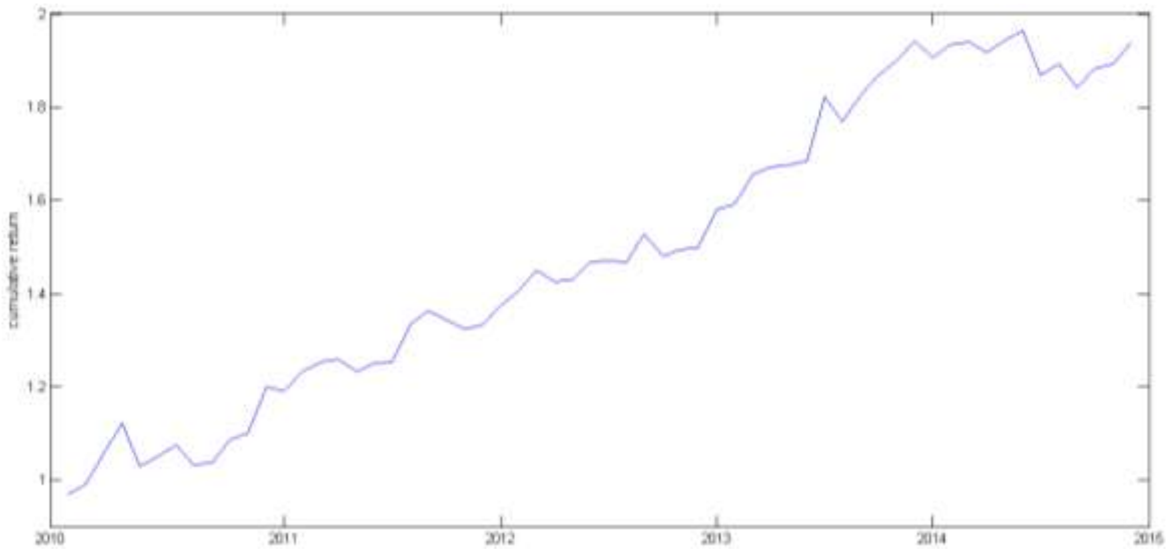


Figure 164   out-of -sample optimal portfolio with capital allocation and risk aversion coefficient = 1 realized returns

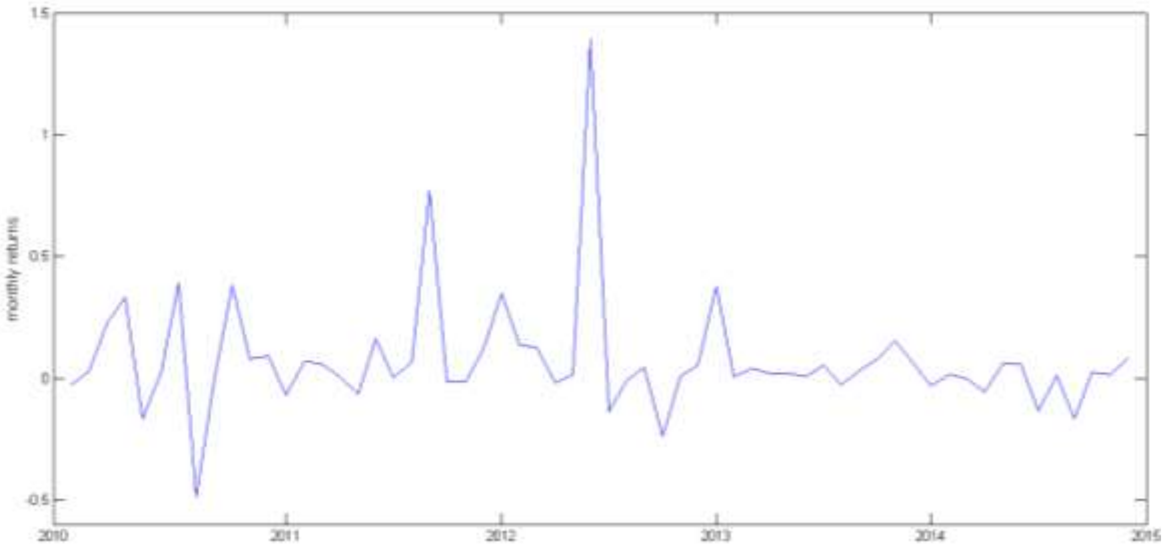


Figure 165 out-sample optimal portfolio with capital allocation and risk aversion coefficient = 1 cumulative realized returns

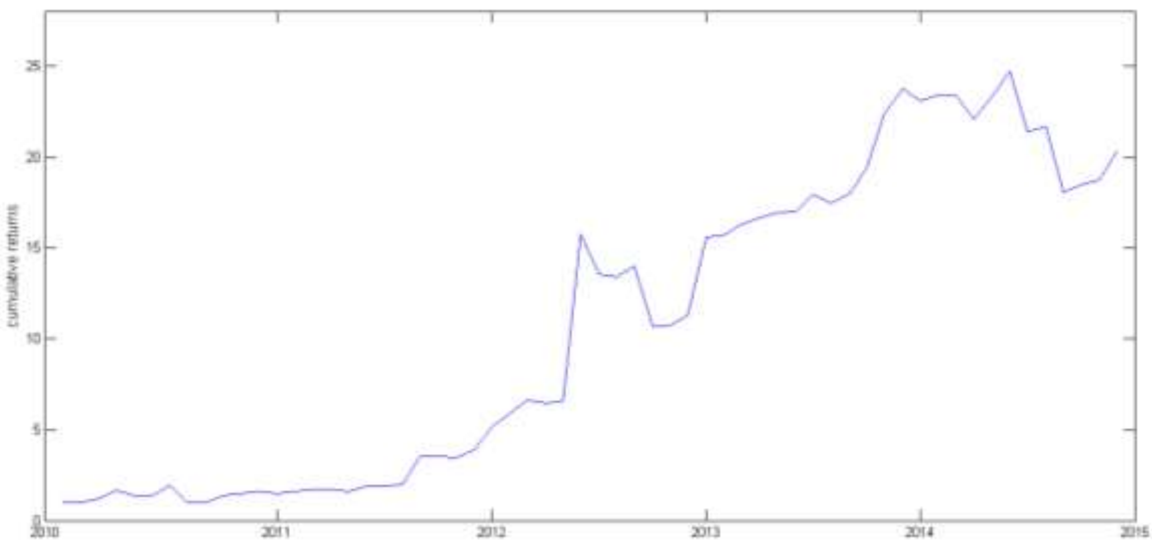


Figure 166 out-of -sample optimal portfolio with capital allocation and risk aversion coefficient = 3 realized returns

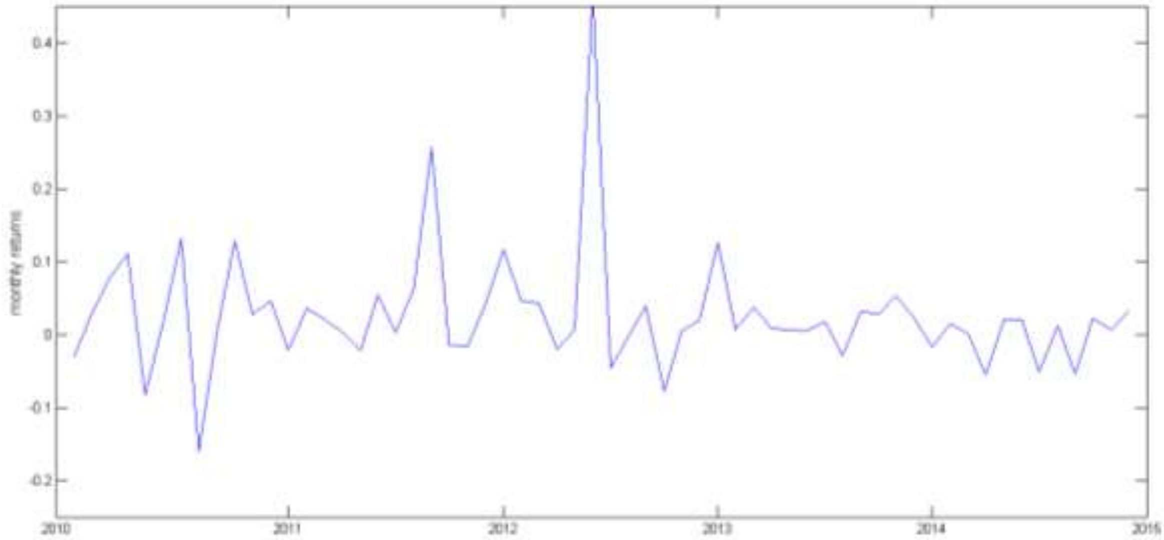


Figure 167 out-of -sample optimal portfolio with capital allocation and risk aversion coefficient = 3 cumulative realized returns

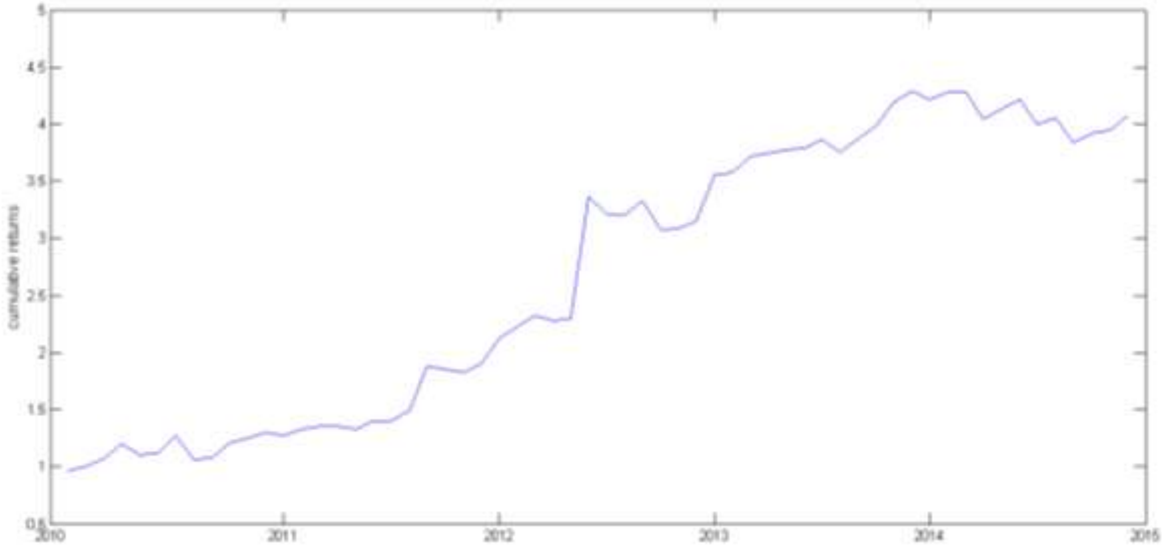


Figure 168 out-of -sample optimal portfolio with capital allocation and risk aversion coefficient = 5 realized returns

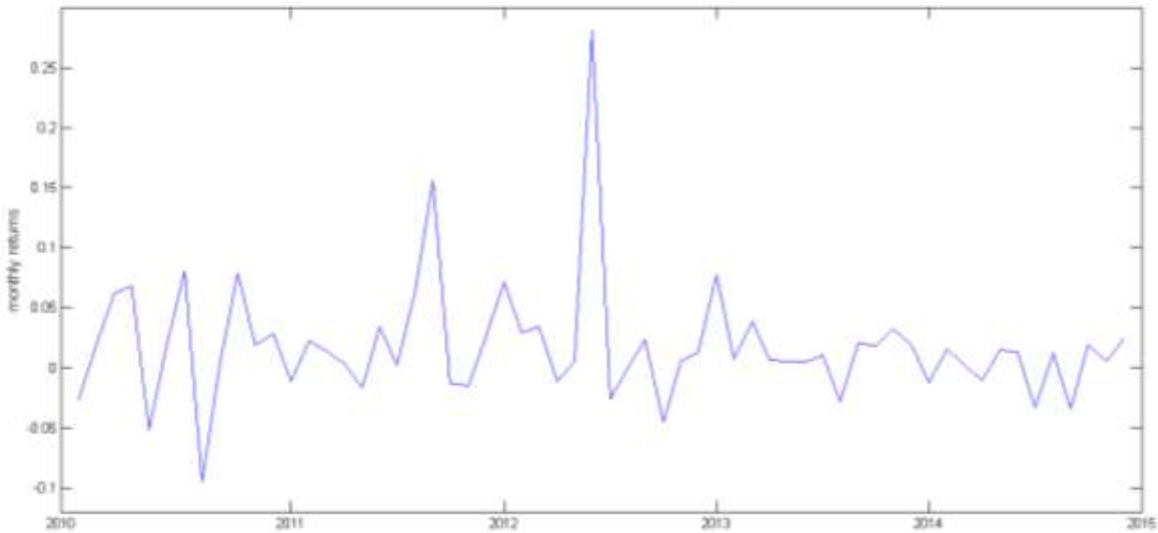
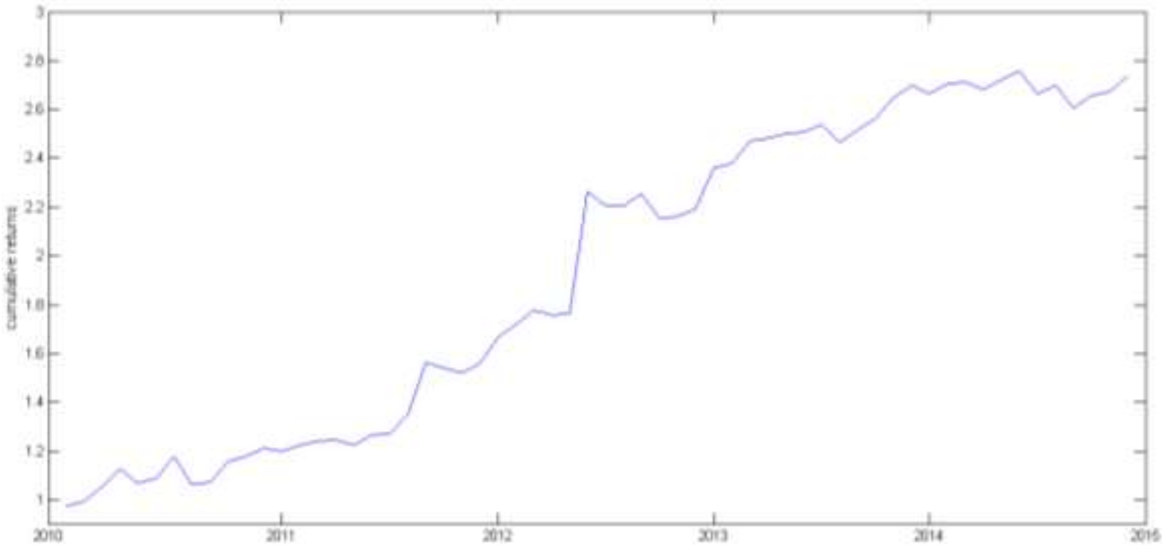


Figure 169 out-of -sample optimal portfolio with capital allocation and risk aversion coefficient = 5 cumulative realized returns



4.3.2 *Out-of-sample asset allocation exercise based on a VAR(1) model*

In this subparagraph the out-of-sample portfolio construction results using the VAR(1) are shown. The VAR(1) model parameters have been estimated using the in-sample returns. The same five out-of-sample recursive optimal portfolios as in the MSVAR(2,1) case have been built. Similarly to what have been done for the MSVAR(2,1) out-of-sample vector of dynamic conditional mean, the one step ahead prediction *(t+1)* uses the current returns *(t)*; it follows that the time *t+T*

prediction, which for example might be the prediction for December 2011 (t=540 that corresponds to December 2009 and T=12) uses the November 2011 (t=540 that corresponds to December 2009 and T=11) realized returns while the predicted variance-covariance matrix for each *t+T* from January 2010 to December 2014 is constant since the VAR(1) variance-covariance is not time varying.

Figure 170 out-of-sample lo20 and its VAR(1) dynamic conditional mean

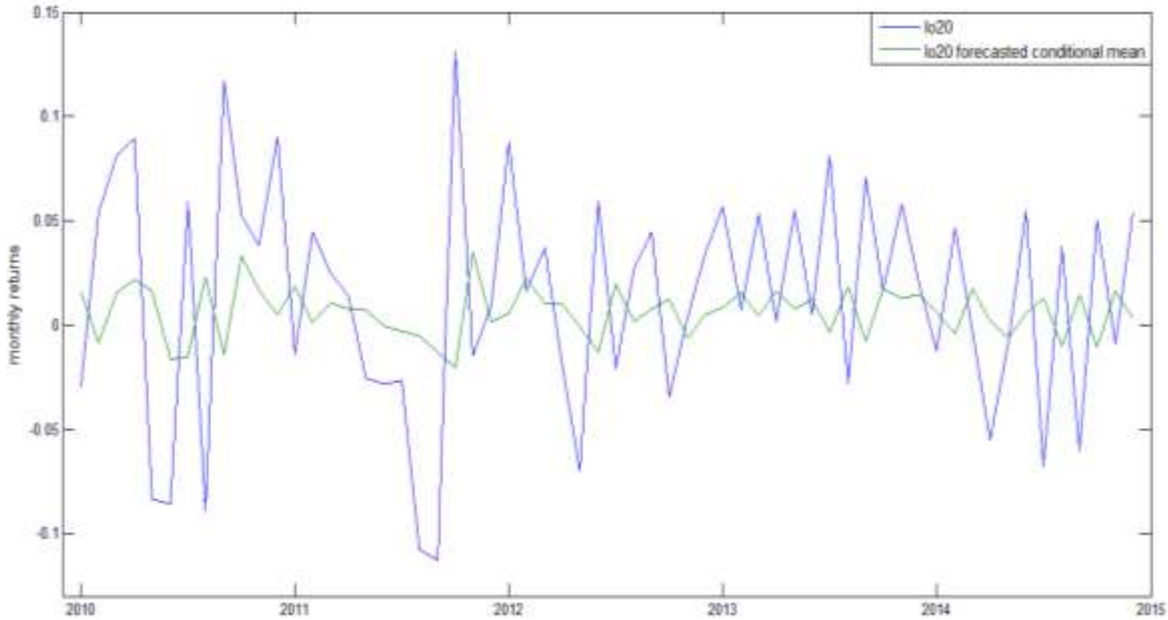


Figure 171 out-of-sample hi20 and its VAR(1) predicted conditional mean

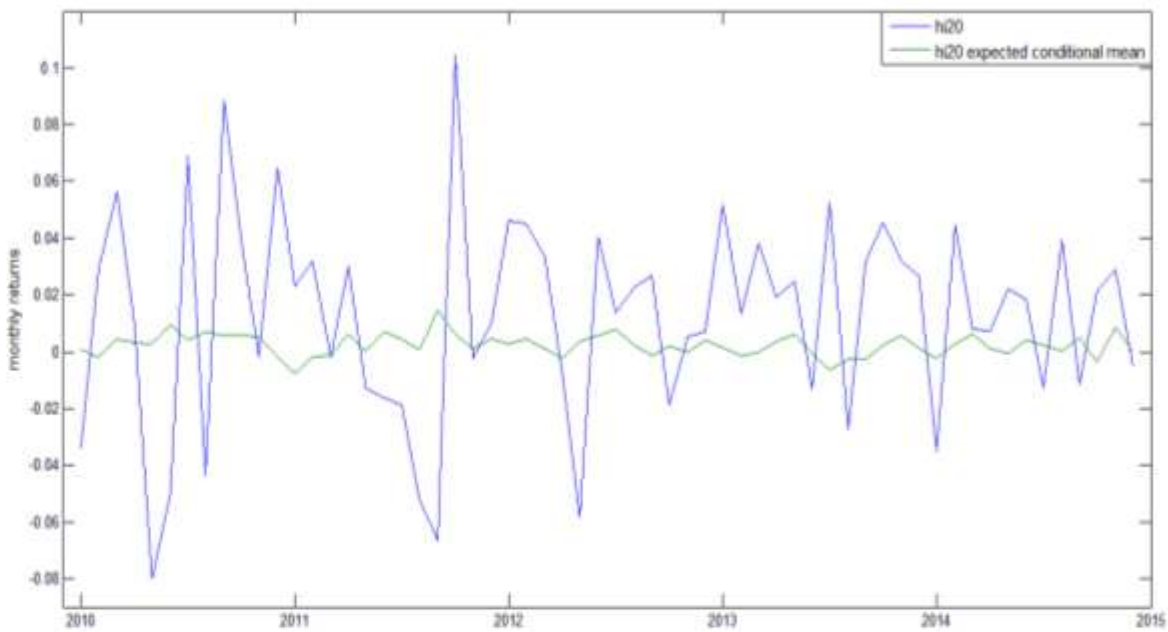


Figure 172 out-of-sample tbond and its VAR(1) predicted conditional mean

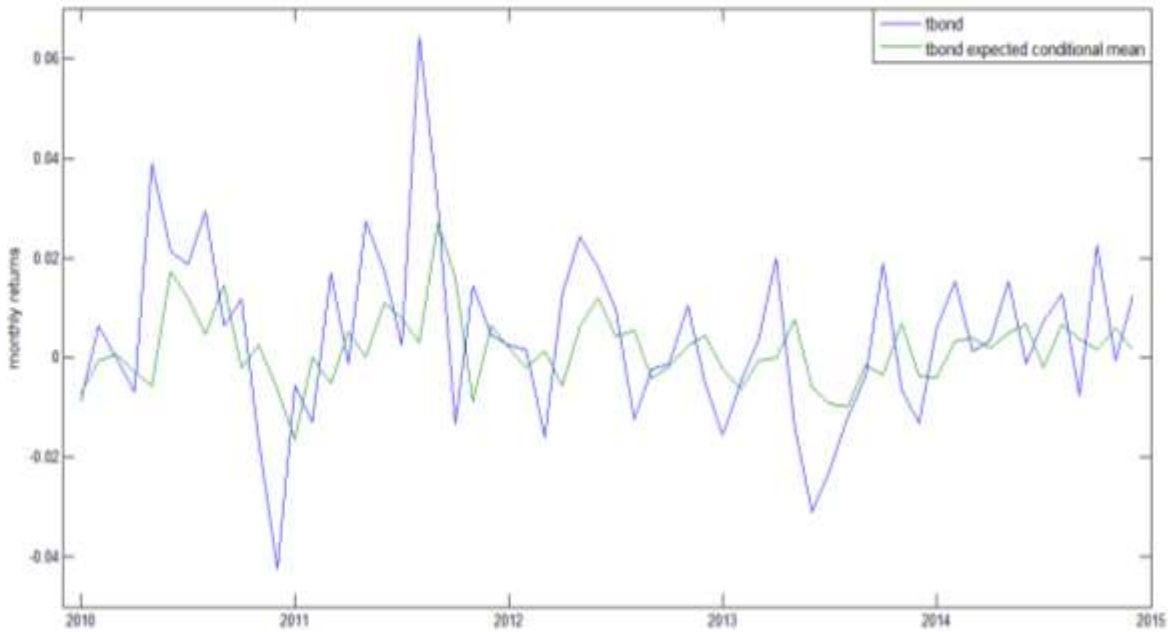


Figure 173 VAR(1) out-of-sample maximum Sharpe ratio portfolios weights (1 = 100%) with opened lower budget constraint (permit to invest in the riskless asset)

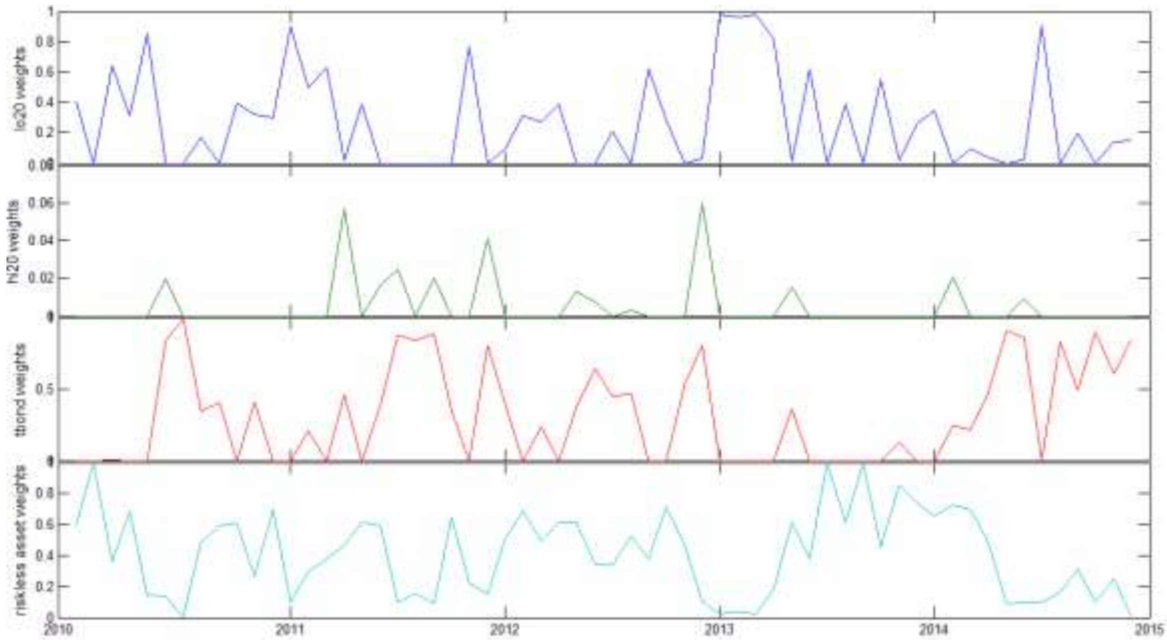


Figure 174 VAR(1) out-of-sample maximum Sharpe ratio portfolios weights (1 = 100%) with budget constraint (not permit to invest in the riskless asset)

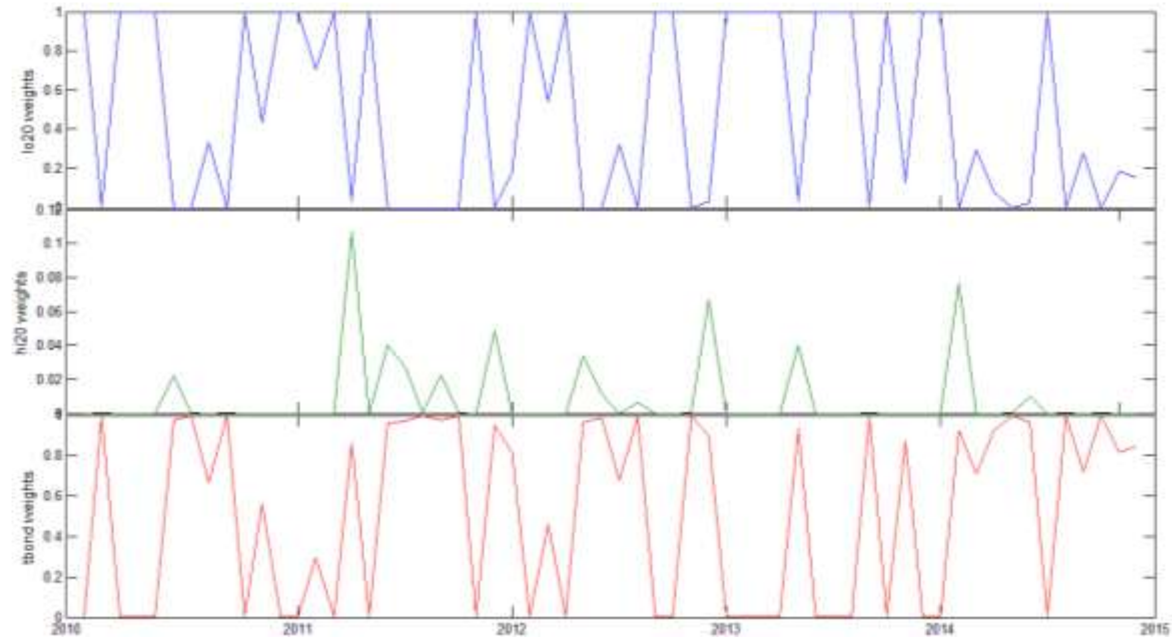


Figure 175 VAR(1) out-of-sample overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 1

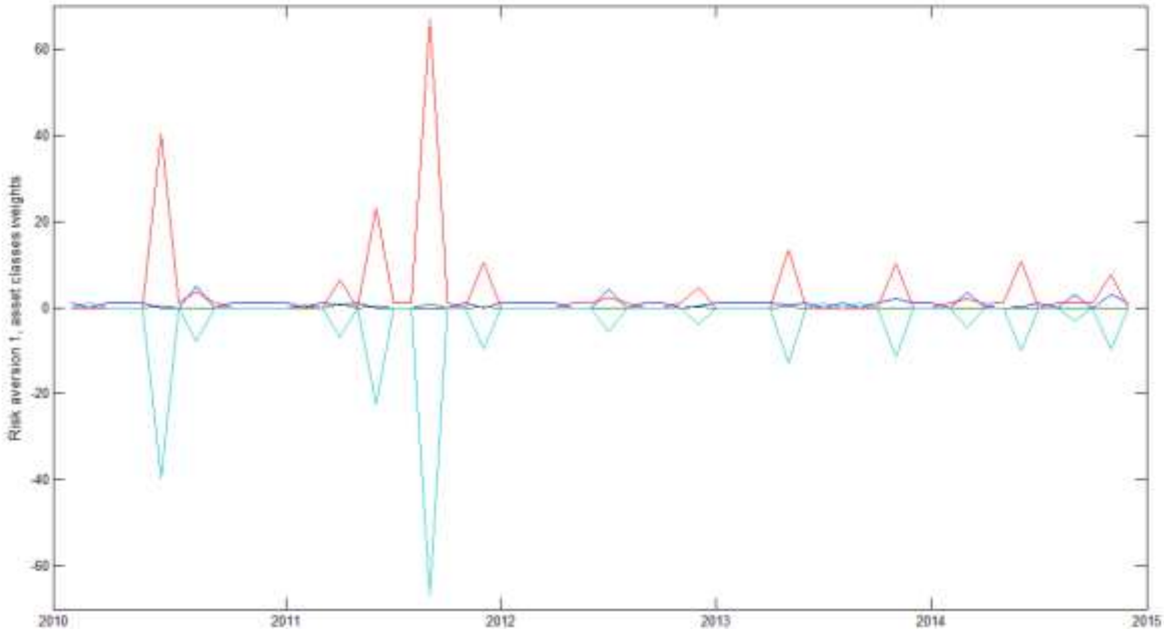


Figure 176 VAR(1) out-of-sample overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 3

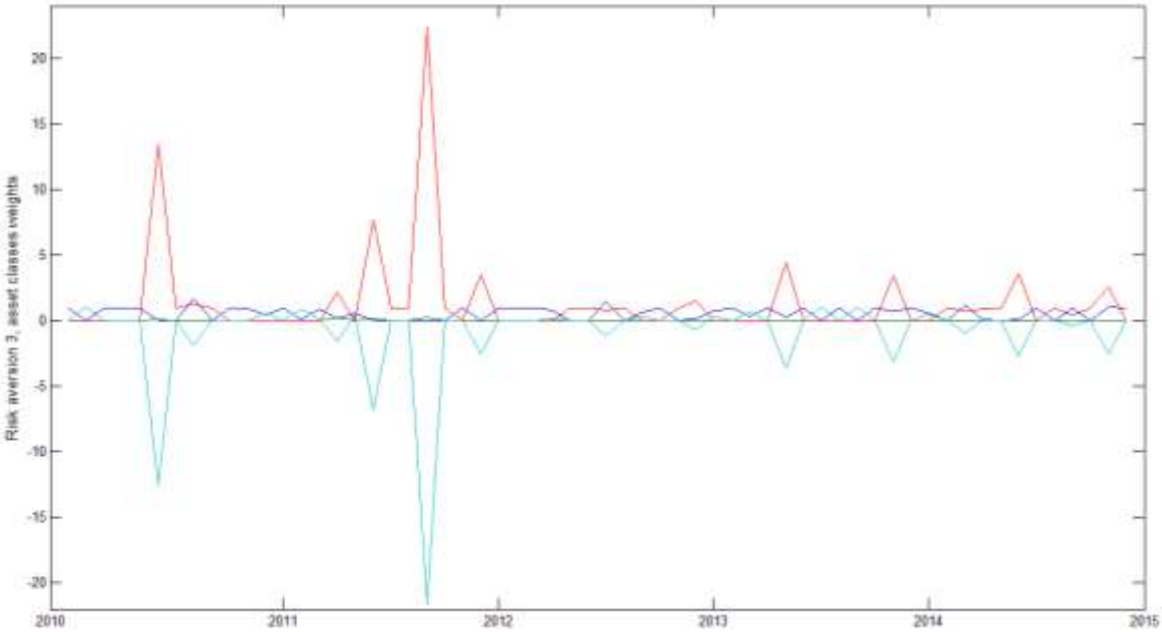


Figure 177 VAR(1) out-of-sample overall optimal portfolio weights (1 = 100%) with capital allocation and risk aversion coefficient = 5

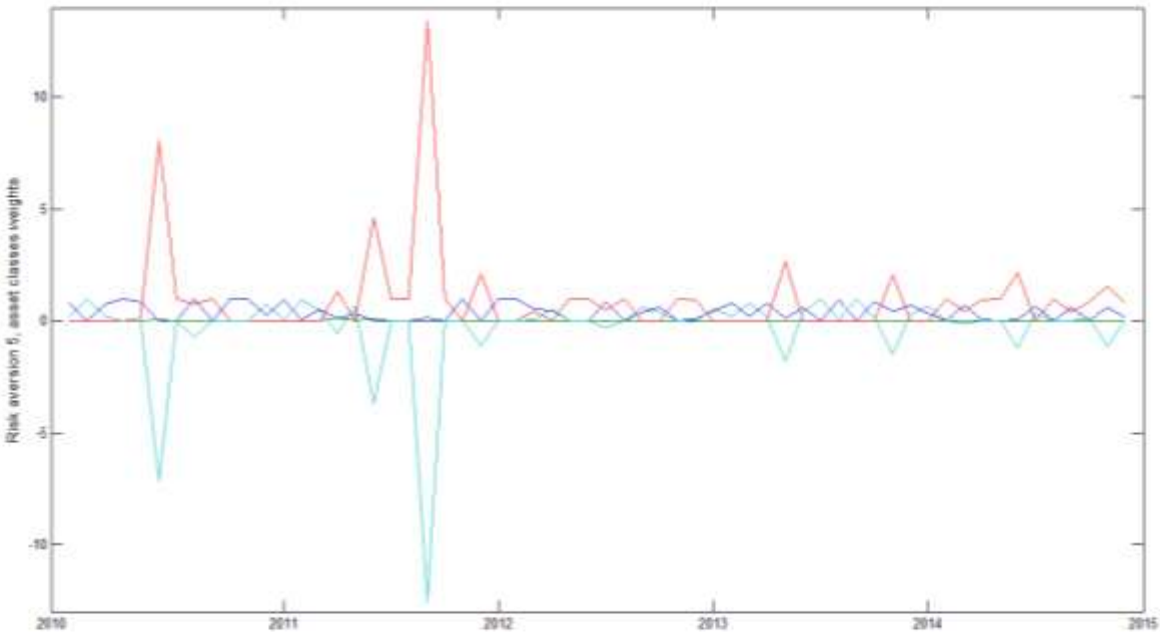


Table 34 VAR(1) out-of-sample average portfolio overall weights (1 = 100%)

avg. weights maximum Sharpe ratio portfolio with opened lower budget constraint			
lo20	*hi20*	*tbond*	*risk-free asset*
0.2709 (0.3148)	0.0051 (0.0128)	0.3107 (0.3411)	0.4132 (0.2747)

avg. weights maximum Sharpe ratio portfolio with budget constraint							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.4788	(0.4640)	0.0085	(0.0210)	0.5126	(0.4557)	0.0000	(0.0000)
avg. optimal weights with capital allocation and risk aversion coefficient = 1							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.8531	(1.0363)	0.0382	(0.1527)	3.6249	(10.5283)	-3.5162	(10.5595)
avg. optimal weights with capital allocation and risk aversion coefficient = 3							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.5457	(0.4914)	0.0127	(0.0509)	1.3633	(3.4813)	-0.9218	(3.3899)
avg. optimal weights with capital allocation and risk aversion coefficient = 5							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.3943	(0.3845)	0.0076	(0.0305)	0.9320	(2.0789)	-0.3339	(1.9918)

Table 35 VAR(1) out-of-sample average portfolio risky weights (1 = 100%)

avg. optimal risky weights with capital allocation and risk aversion coefficient = 1							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.5355	(0.4712)	0.0023	(0.0124)	0.4123	(0.4621)	0.0000	(0.0000)
avg. optimal risky weights with capital allocation and risk aversion coefficient = 3							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.5340	(0.4725)	0.0023	(0.0124)	0.4137	(0.4636)	0.0000	(0.0000)
avg. optimal risky weights with capital allocation and risk aversion coefficient = 5							
lo20		*hi20*		*tbond*		*risk-free asset*	
0.5098	(0.4663)	0.0023	(0.0124)	0.4380	(0.4602)	0.0000	(0.0000)

Figure 178 VAR(1) out-of-sample maximum Sharpe ratio portfolios (with and without budget constraint) expected returns

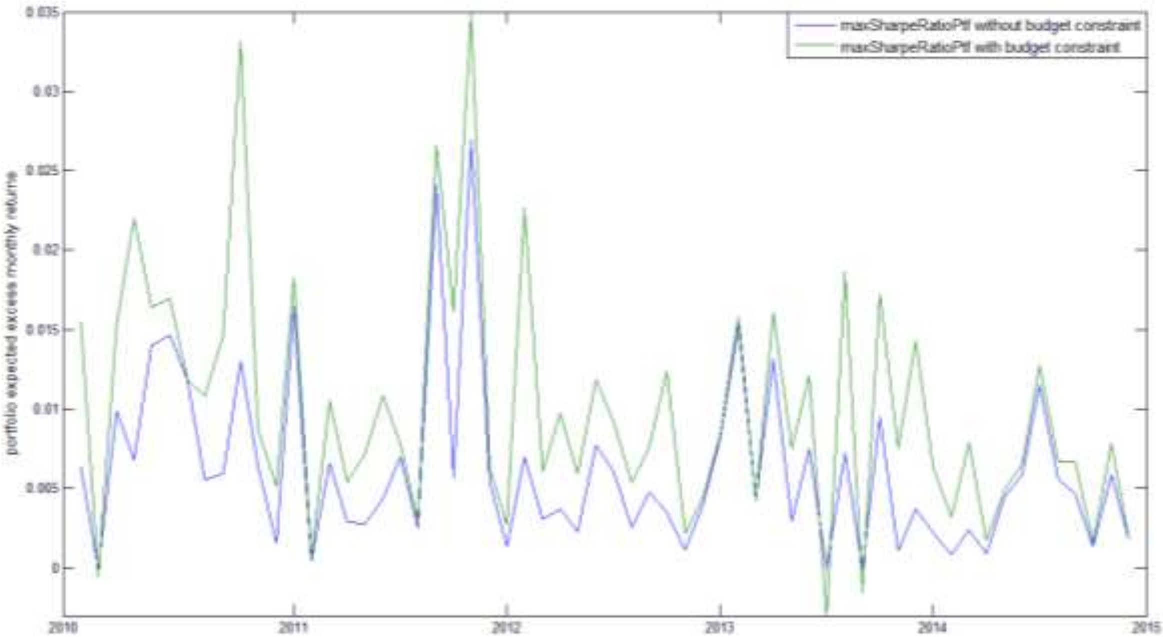


Figure 179 VAR(1) out-of-sample expected returns of the maximum expected utility portfolios (capped at 1 = 100%)

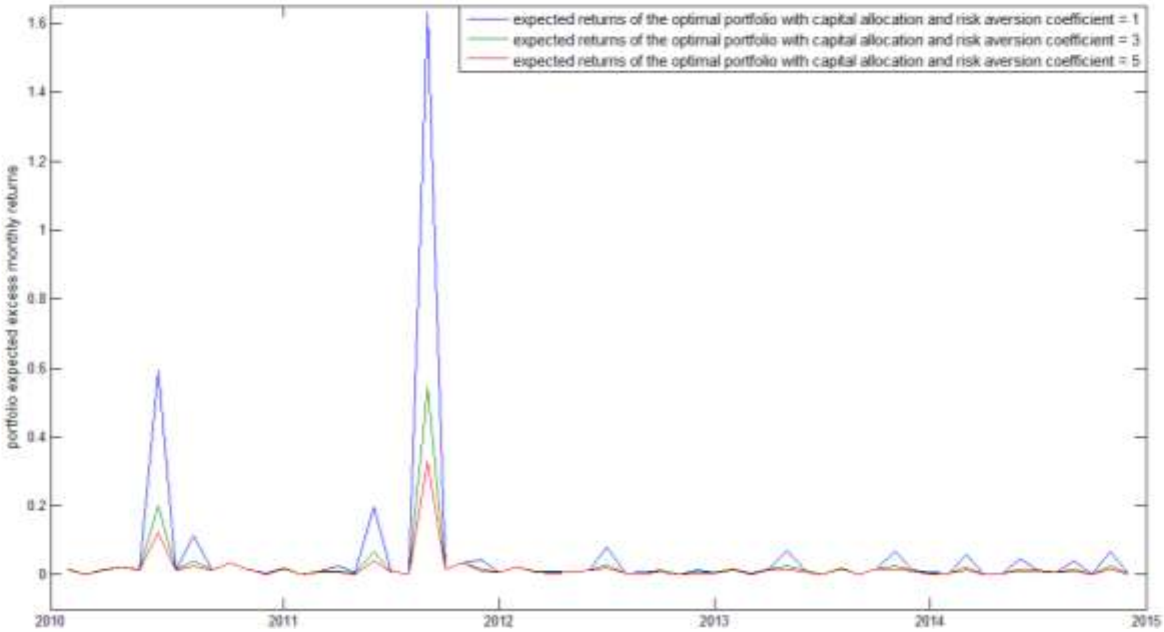


Table 36 VAR(1) out-of-sample portfolios expected moments

maximum Sharpe ratio portfolio with opened lower budget constraint		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0061 (0.0055)	0.0222 (0.0165)	0.3034 (0.2459)
maximum Sharpe ratio portfolio with budget constraint		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*

0.0101 (0.0077)	0.0386 (0.0212)	0.3007 (0.2496)
optimal weights with capital allocation and risk aversion coefficient = 1		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0588 (0.2214)	0.1163 (0.1976)	0.2747 (0.2288)
optimal weights with capital allocation and risk aversion coefficient = 3		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0251 (0.0733)	0.0581 (0.0622)	0.2800 (0.2285)
optimal weights with capital allocation and risk aversion coefficient = 5		
avg. expected return	*avg. expected standard deviaton*	*avg. Sharpe ratio*
0.0175 (0.0441)	0.0410 (0.0376)	0.2860 (0.2282)

Table 37 VAR(1) out-of-sample realized portfolios statistics

maximum Sharpe ratio portfolio with opened lower budget constraint		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0054	0.0210	0.2560
maximum Sharpe ratio portfolio with budget constraint		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0101	0.0325	0.3093
optimal weights with capital allocation and risk aversion coefficient = 1		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0438	0.2651	0.1651
optimal weights with capital allocation and risk aversion coefficient = 3		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0203	0.0922	0.2204
optimal weights with capital allocation and risk aversion coefficient = 5		
average realized return	*realized returns standard deviaton*	*realized returns Sharpe ratio*
0.0146	0.0581	0.2518

Figure 180 VAR(1) out-of-sample maximum Sharpe ratio portfolio with opened lower budget constraint realized returns

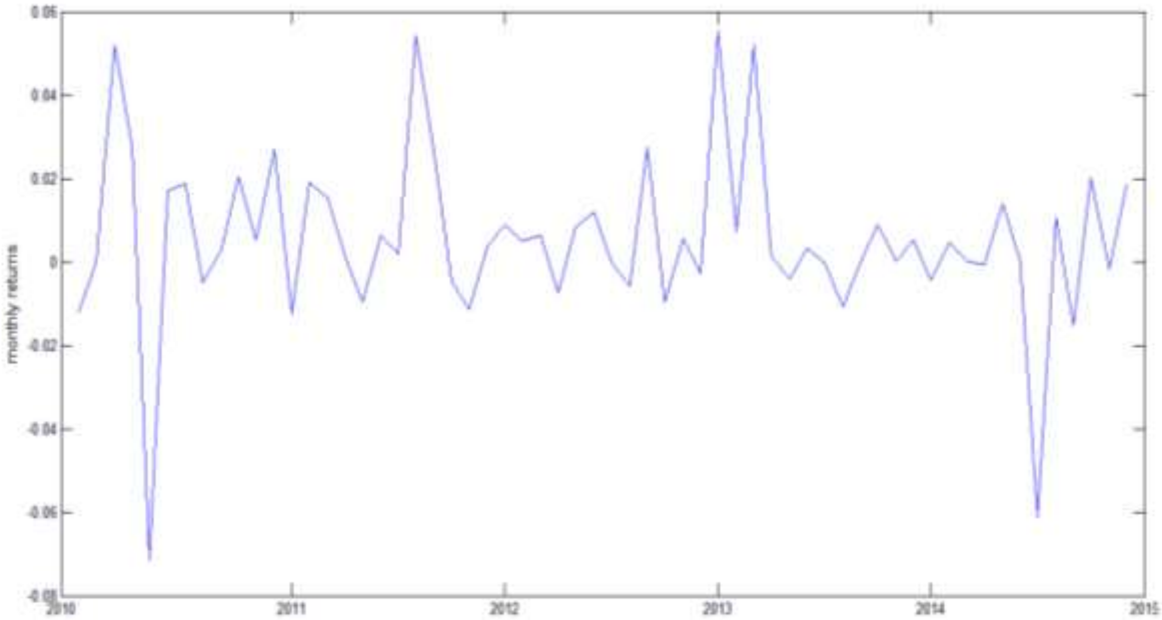


Figure 181 VAR(1) out-of -sample maximum Sharpe ratio portfolio with opened lower budget constraint cumulative realized returns

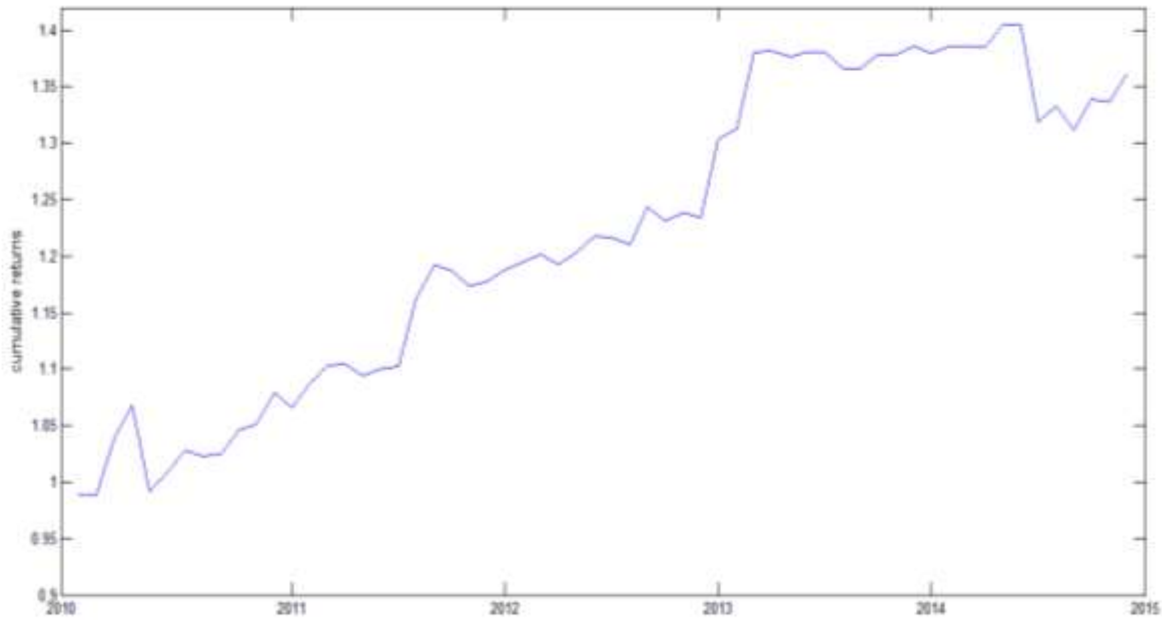


Figure 182 VAR(1) out-of -sample maximum Sharpe ratio portfolio with budget constraint realized returns

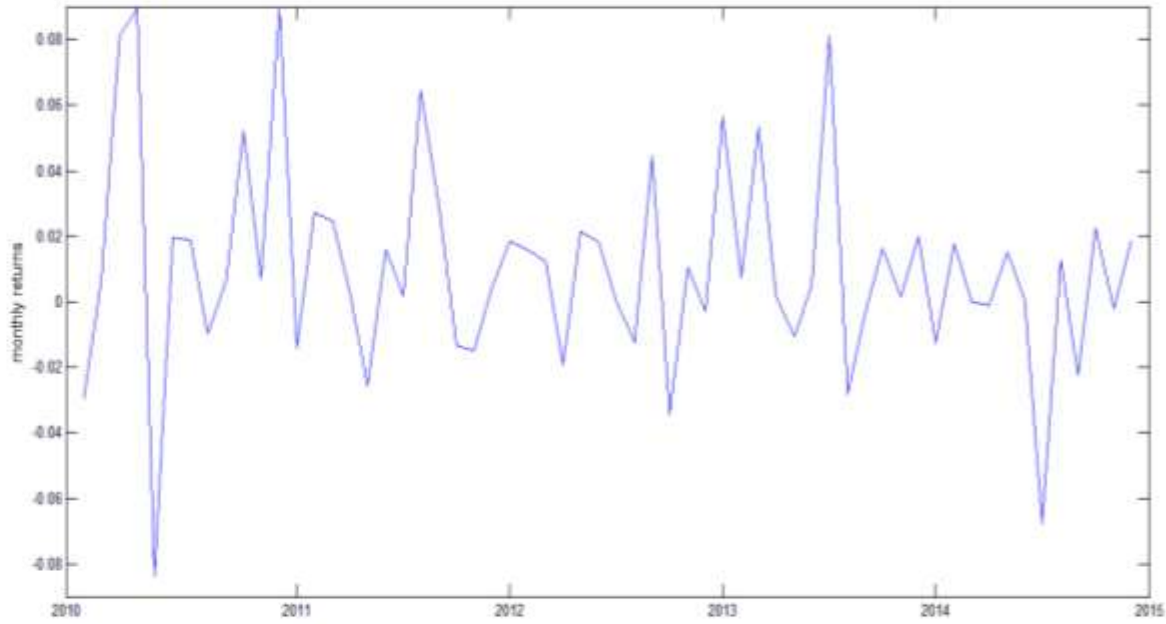


Figure 183 VAR(1) out-of -sample maximum Sharpe ratio portfolio with budget constraint realized cumulative returns

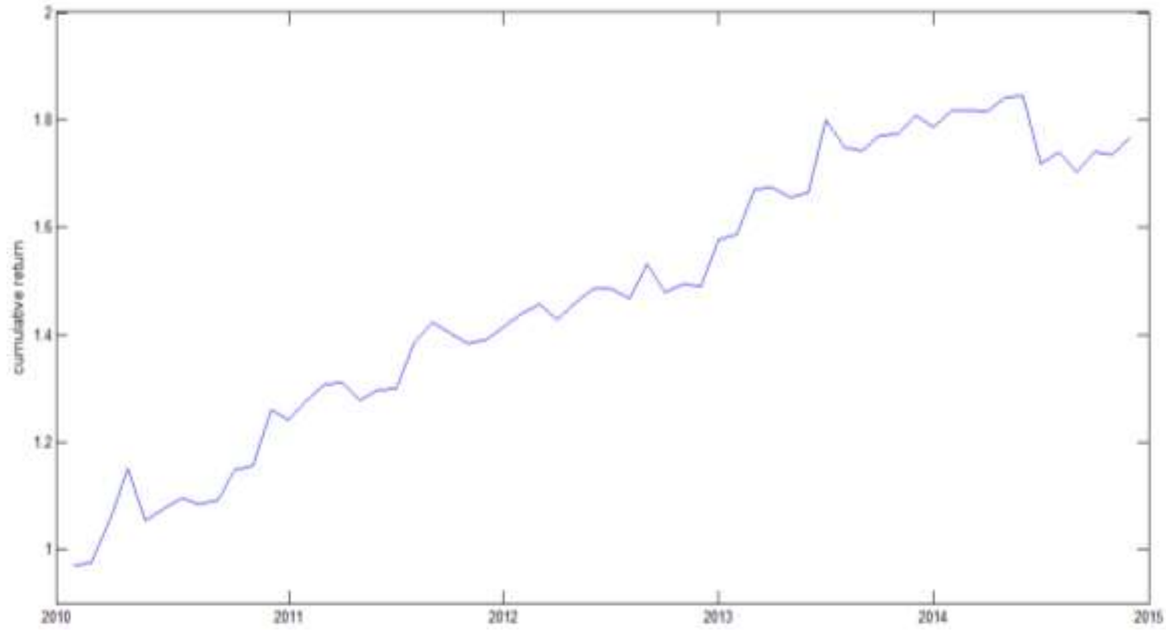


Figure 184 VAR(1) out-of -sample optimal portfolio with capital allocation and risk aversion coefficient = 1 realized returns

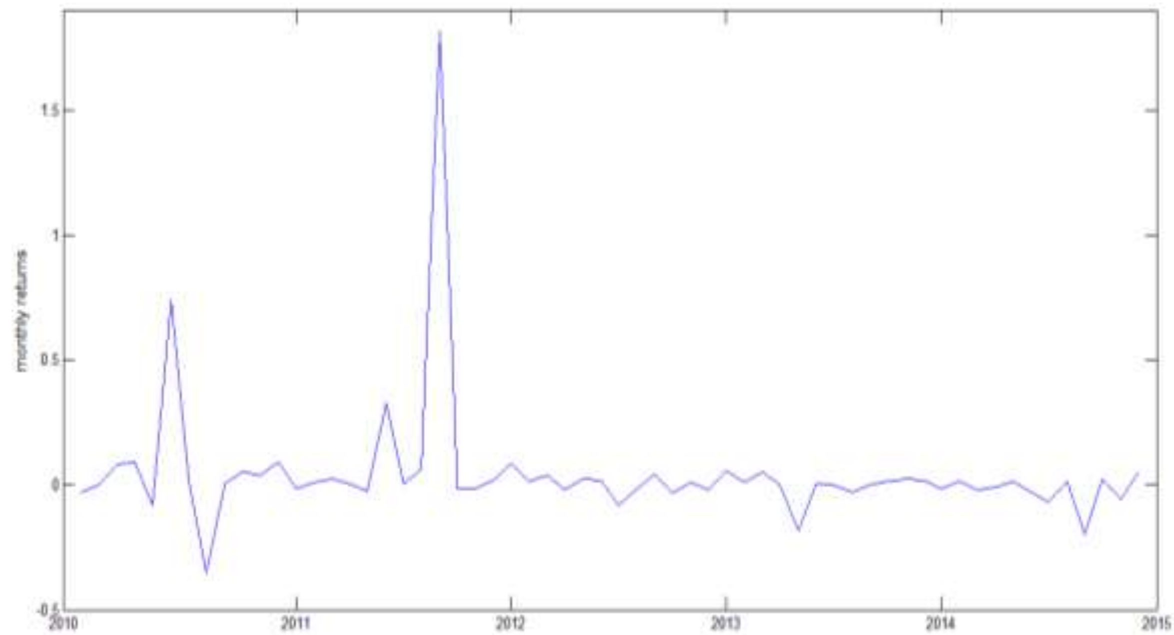


Figure 185 VAR(1) out-sample optimal portfolio with capital allocation and risk aversion coefficient = 1 cumulative realized returns

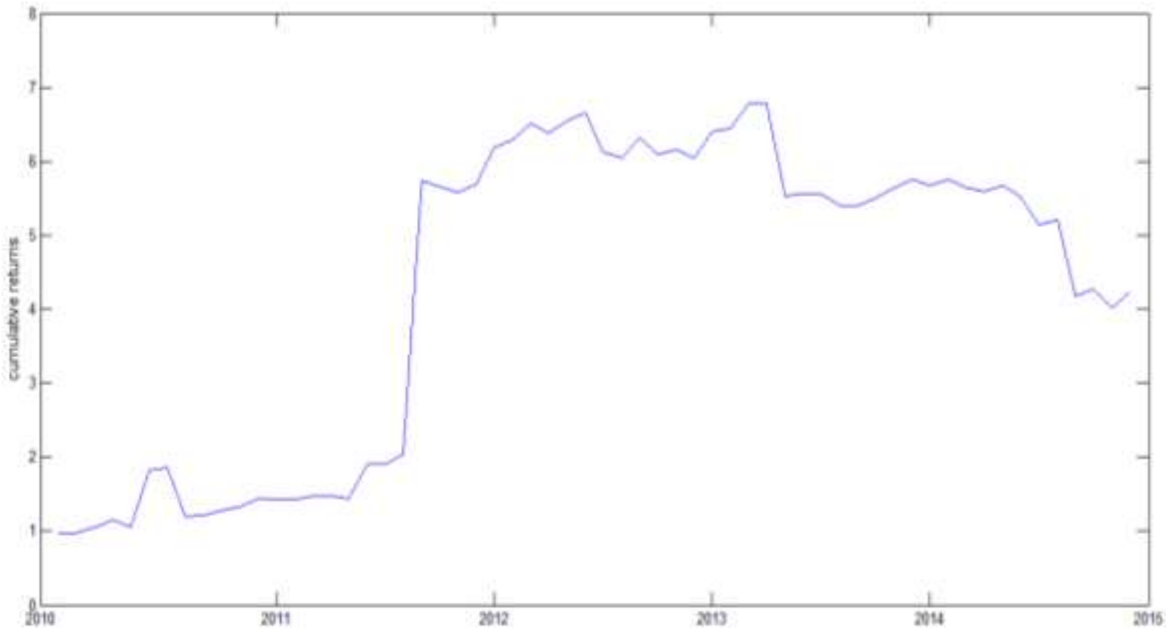


Figure 186 VAR(1) out-of -sample optimal portfolio with capital allocation and risk aversion coefficient = 3 realized returns

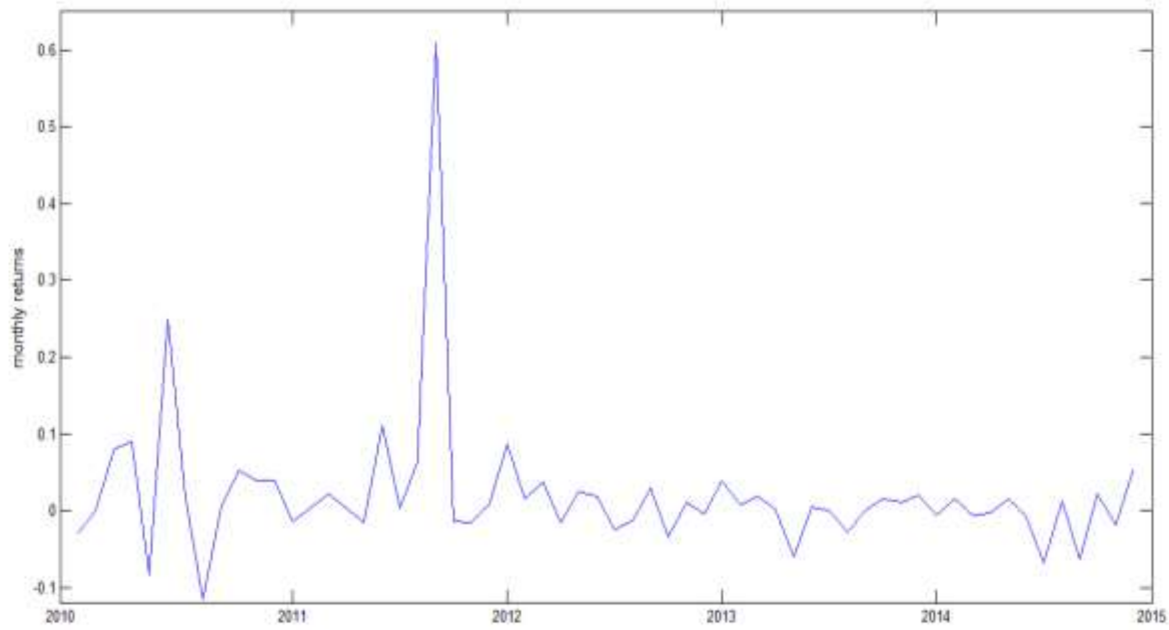


Figure 187 VAR(1) out-of -sample optimal portfolio with capital allocation and risk aversion coefficient = 3 cumulative realized returns

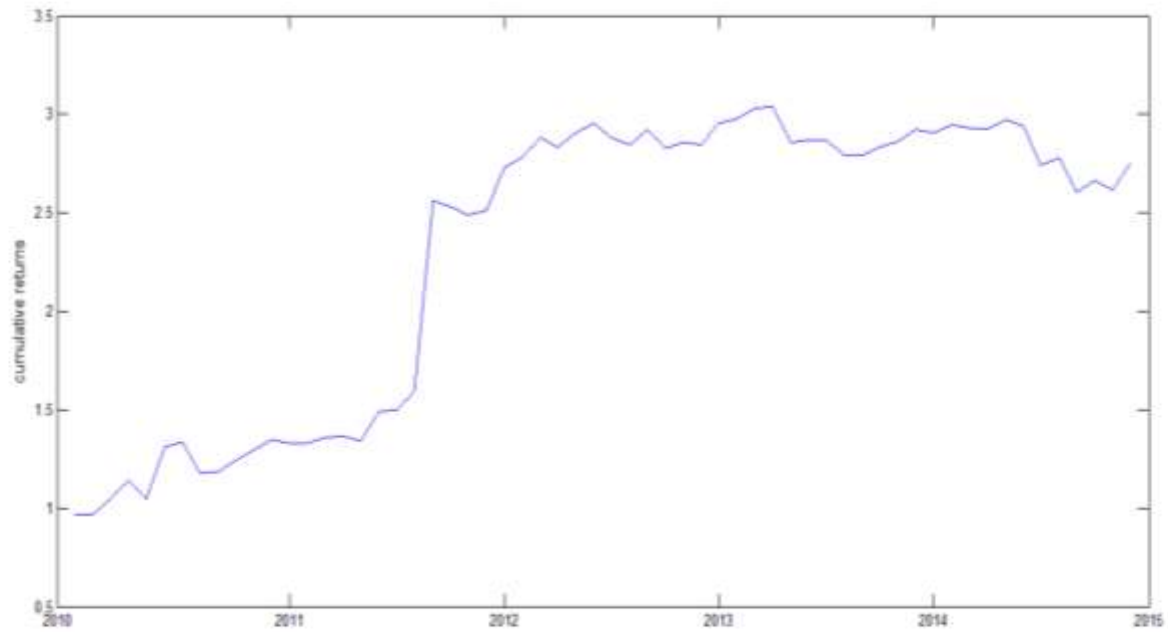


Figure 188 VAR(1) out-of -sample optimal portfolio with capital allocation and risk aversion coefficient = 5 realized returns

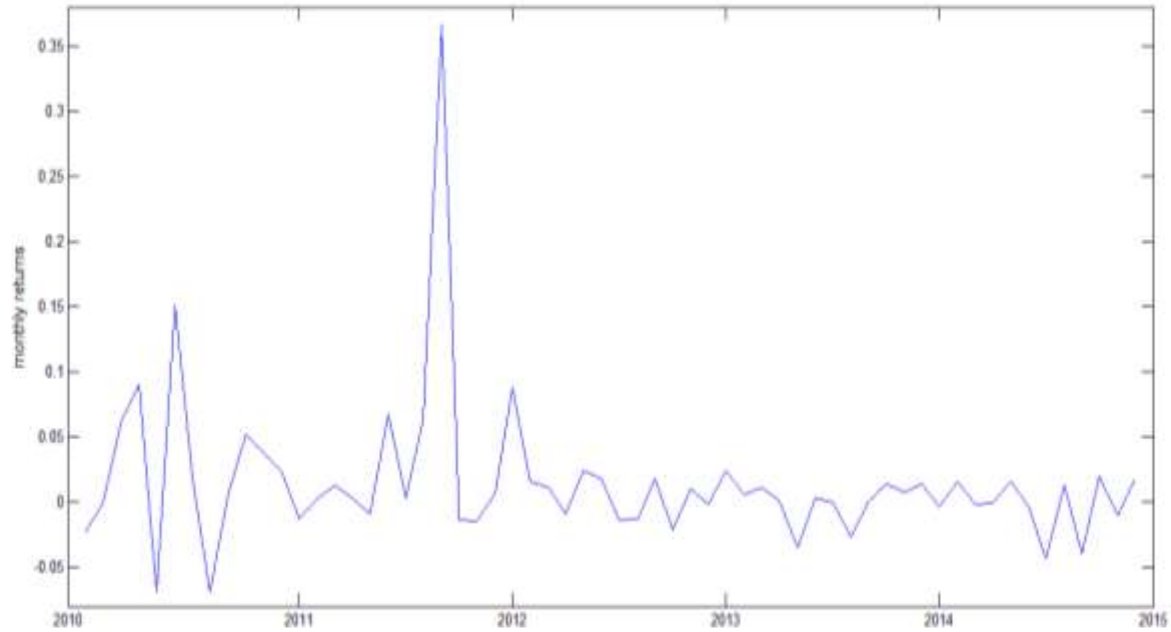
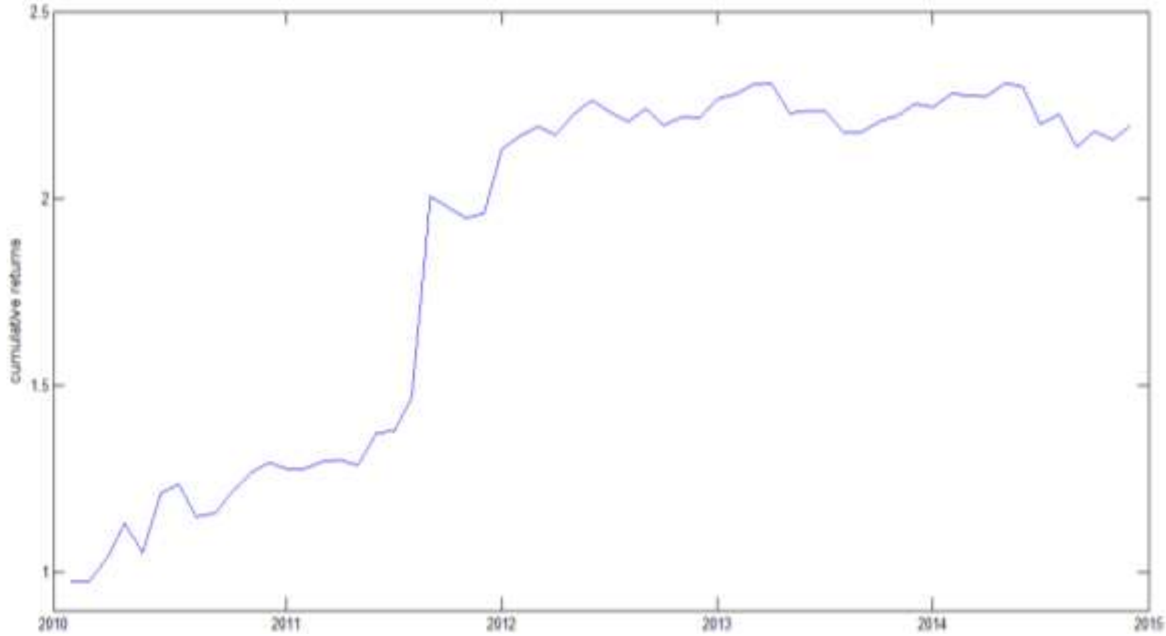


Figure 189 VAR(1) out-of -sample optimal portfolio with capital allocation and risk aversion = 5 cumulative realized returns



Conversely to the in the in-sample, in the out-of-sample exercise the portfolios based on the MSVAR(2,1) asset classes first two moments estimates did largely overperform those built on the VAR(1) model. What emerges from a comparison of Table 33 and Table 37 is that in all the five types of portfolios the MSVAR(2,1)

leads to greater average portfolio realized returns and to greater realized Sharpe ratios compared to the VAR(1).

4.4 Forecasting ability comparison

In this paragraph an assessment of the in-sample and the out-of-sample one-step-ahead forecasting ability of both the single regime VAR(1) and the regime switching MSVAR(2,1) model is illustrated and discussed. Three different criteria have been adopted in this comparison analysis.

4.4.1 RMSE

The one-step-ahead average root-mean squared error is given by the formula:

$$(61) \quad RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\boldsymbol{y}_{\mathbf{t+1}} - \mathrm{E}(\boldsymbol{y}_{\mathbf{t+1}})\right)^2}$$

where $\boldsymbol{y}_{\mathbf{t+1}}$ are the actual small stock, large stocks and bond realized returns while $\mathrm{E}(\boldsymbol{y}_{\mathbf{t+1}})$ represents the forecasted returns; both vectors are constructed as a sequence of the single asset classes returns, thus if the length of the sample is 60 periods and the model has three asset classes (as in the out-of-sample exercise) the length of either $\boldsymbol{y}_{\mathbf{t+1}}$ or $\mathrm{E}(\boldsymbol{y}_{\mathbf{t+1}})$ is 180. The RMSE is a frequently used measure of the differences between values predicted by a model and the values actually observed; it represents the sample standard deviation of the differences between predicted values and actual realized returns. The RMSE is a good measure of accuracy to compare forecasting errors of different models, thus smaller RMSE are preferred.

Table 38 in-sample and out-of-sample RMSE comparison

VAR(1)	
In-sample RMSE	***Out-of-sample RMSE***
0.0387	0.0416
MSVAR(2,1)	
In-sample RMSE	***Out-of-sample RMSE***
0.0451	0.0411

Table 38 shows both the in-sample and the out-of-sample VAR(1) and MSVAR(2,1) RMSE values in order to assess their forecasting ability. As it can be seen the simple VAR(1) model performs better than the more complicated MSVAR(2,1) model in the in-sample period but there is no evidence of this relative advantage in the out-of-sample period where the two models substantially have the same forecasting ability.

4.4.2 Pearson correlation coefficient

The second measure I adopted is the Pearson correlation coefficient between the forecasted return and the actual realized ones. The closer the slope coefficient to 1, the better the forecasting ability of the model.

Table 39 in-sample and out-of-sample Pearson correlation coefficient between realized and forecasted returns by the VAR(1) and the MSVAR(2,1)

VAR(1)	
In-sample correlation	***Out-of-sample correlation***
0.7816	-0.0751
MSVAR(2,1)	
In-sample correlation	***Out-of-sample correlation***
0.2590	-0.0048

In the in-sample period both models seem to forecast returns in the right direction, however the correlation coefficient is smaller in absolute value for the regime switching model. In the out-of-sample period both models seem to have very weak but negative predictive power, which means that both models forecast returns in the wrong direction. This finding substantially confirms the conclusion based on the RMSE comparison.

4.4.3 Regression analysis

As a third measure I regress the realized returns for each risk factor on the forecasted returns and an intercept.

(62) $$\boldsymbol{y}_{\mathbf{t+1}} = a + b\mathrm{E}(\boldsymbol{y}_{\mathbf{t+1}})$$

If the forecasts are accurate, I expect $a = 0$ and $b = 1$.

Table 40 results of the in-sample and out-of-sample regression of the realized returns on the returns forecasted by the VAR(1) and the MSVAR(2,1)

VAR(1)		
	Intercept	***Slope***
In-sample	-0.0093	3.3181
Out-of-sample	0.0110	-0.3402
MSVAR(2,1)		
	Intercept	***Slope***
In-sample	-0.0005	0.9363
Out-of-sample	0.0100	-0.0182

A comparison of the intercept for the regime switching model and the single state model reveals that in the in-sample period both models underestimate future returns, while in the out-of-sample period they overestimate them. The slope coefficients indicate that the relationships between forecasted returns and realized returns in the in-sample period is stronger in the MSVAR(2,1) model than in the VAR(1) model while in the out-of-sample period there is no substantial difference between the two models.

4.5 Asset allocation exercise conclusions

In conclusion it can be said that while the construction of portfolios based on the two regimes MSVAR(2,1) model estimates do not lead to results that clearly overperform those based on the single regime VAR(1) model estimates in the in-sample framework, it significantly does it in the out-of-sample framework. Therefore the adoption of a sophisticated multi regimes model greatly improves the portfolios performance in the out-of-sample exercise for all types of portfolios constructed.

CHAPTER 5 - ASSET ALLOCATION BASED ON THE COPULA-OPINION POOLING APPROACH

The purpose of this paragraph is to investigate and make a comparison of the behavior of the MSVAR and the VAR in the COP framework in terms of both expected and realized portfolio returns and Sharpe ratio in the context of mean-variance and conditional value-at-risk (CVaR) portfolio optimization. In the first part of the each subparagraph the results of an out-of-sample portfolio optimization exercise based on the copula opinion pooling approach applied both to the simple single regime VAR(1) model and to the two regimes MSVAR(2,1) model are shown; in the second part of each subparagraph some conclusions regarding the performances of the portfolios based on the two models are drawn.

5.1 The copula-opinion pooling approach

The recent financial crisis has pointed out the “Risk-on/Risk-off” notion, characterized by periods when the market is not driven any more by the fundamental analysis. Consequently, a question about the effectiveness of the classic portfolio optimization models has risen: the strong rise of cross asset correlations suggests integrating new market anomalies into models, such as the non-normality of returns, the non-linear dependence, the sharp rise of correlations between assets historically uncorrelated that occurred during the recent recession or the fact that the linear codependence modeled by a variance-covariance matrix in an asset returns normality framework are not well adapted for the stressed scenarios that have been observed over the out-of-sample years, namely from January 2010 to December 2014.

The Markov regime switching models characteristically fit these returns features given their non-linear nature; but what is further needed is a more complex portfolio optimization approach that can be able to take into account and to manage these recently observed asset returns features.

As a consequence it seems necessary to widen the possibilities of the Black-Littermann model to take these anomalies into account. As Meucci (2005) reminded, the pathbreaking technique by Black and Litterman (1990) (BL) allows

portfolio managers to compute a posterior market distribution that smoothly blends their subjective views on the market with a prior market distribution, under the assumption that all the distributions involved be normal. It is based on both Bayesian markets forecasts and mean-variance optimization, it solves the issue of stability on expected returns, and proposes a way to rely both on fundamental and quantitative analysis. In reality, as already established, the markets are in general highly non-normal. Furthermore, practitioners might wish to input their views in less informative ways than the "alpha + normal noise" prescription of BL, for instance by means of uniform distributions on ranges.

In addition, the COP approach allows the investor to set views on the market instead of the market parameters.
In theory, these issues are fully addressed by the copula-opinion pooling (COP) approach, see Meucci (2006).
It turns out that the implementation in practice of the COP approach under any distributional assumption is straightforward: by representing the prior market distribution in terms of a very large number of Monte Carlo scenarios, an equal number of scenarios from the COP posterior market distribution can be computed in a fraction of a second. Also, Monte Carlo methods are not cursed by dimensionality: therefore the COP approach can handle markets with any number of risk factors.

The intuition behind the COP approach is to determine the marginal distribution of each view separately, whereas the joint co-dependence, i.e. the copula, among the views is directly inherited from the prior market structure. Subsequently the joint distribution of the views is translated into a joint posterior distribution for the market. Thus, the first objective of the model is to simulate the expected future evolution of the market, the so called posterior market distribution, mixing the results of a quantitative model based on historical data, the so called prior market distribution, and the views of the investor. These two cumulative distribution functions are then averaged according to weights depending on the confidence levels the investor has on his views.

The representation of the prior distribution depends on two steps, namely modeling and simulation. The former one has been performed by the estimation of two alternative models, i.e. the VAR(1) and the MSVAR(2,1). In general the investor can estimate the prior market distribution by means of any techniques and models.

It can be easily said that the in-sample 540 data points (from January 1965 to December 2009) are not enough to motivate the use of empirical distribution. Meucci (2006) suggested parametric bootstrap where the data matrix is first fitted into a multivariate or univariate model, which is later used to simulate another market matrix large enough to motivate the use of empirical distribution. As a consequence the prior market distribution has been obtained by J = 500000 Monte Carlo simulations generating the market prior distribution M of size (*J x N)* where N is equal to the number of time series of the model the investor is interested to invest in, i.e. small stocks, large stocks and bonds. The prior distributions for both VAR(1) and MSVAR(2,1) models have been only created once and stored for further calculations. Thus M represents the returns on the three asset classes. As explained by Meucci (2006) I represent these returns in terms of their multivariate probability distribution function (pdf):

(63) $\mathbf{M} \sim f_{\mathrm{M}}$

I subsequently depart from these distributions according to the subjective views produced by the either the MSVAR(2,1) or the VAR(1) model.

As already explained I empirically represent the prior market distribution by means of a large number of Monte Carlo simulations. More precisely, the prior market distribution is represented by a (*J* × *N)* panel *M*:

(64) $\mathbf{M} \sim f_{\mathrm{M}} \Leftrightarrow \mathcal{M}$

Each N-dimensional row represents one of the J simulated joint scenarios for the market variable M.

The second input of the COP approach are the practitioner's views. Any parametric or non-parametric representation of the cdf of the views is a viable

option. Meucci in his works remarks that in the COP approach this distribution can take on any shape: from the "alpha + normal noise" specification as in BL to an uninformative uniform specification. Also, in order to complete the specification of the views it is needed to set the confidence levels in each view. It is also clear that the views have a huge impact on the final allocation, more or less depending on the confidence level.

I defined uniformly distributed views over the given ranges which can be modeled by the following uniform distribution:

$$(65)\quad \hat{F}_k(v) = \begin{cases} 0 & v \le a_k \\ \frac{v-a_k}{b_k-a_k} & v \in [a_k, b_k], \qquad k = 1,2,3,4 \\ 1 & v \ge b_k \end{cases}$$

This choice probably is not the best one but it is the most general and uninformative one. In my case, for any time over the out-of-sample period, the boundaries of the ranges read:

$$(66)\quad \mathbf{a}' \equiv \begin{pmatrix} \mu_{lo20} - 0.5 * \sigma_{lo20} \\ \mu_{hi20} - 0.5 * \sigma_{hi20} \\ \mu_{tbond} - 0.5 * \sigma_{tbond} \end{pmatrix}$$

$$(67)\quad \mathbf{b}' \equiv \begin{pmatrix} \mu_{lo20} + 0.5 * \sigma_{lo20} \\ \mu_{hi20} + 0.5 * \sigma_{hi20} \\ \mu_{tbond} + 0.5 * \sigma_{tbond} \end{pmatrix}$$

where μ and σ respectively represent at any data point the asset class expected return and standard deviation produced by either the MSVAR(2,1) model or the VAR(1) model . Thus, each time over the out-of-sample period is characterized by a different time varying views since both the MSVAR(2,1) model and the VAR(1) model produce time varying asset class expected returns and standard deviations over time. As a consequence, given the views based on any model, during the out-of-sample period a different returns posterior distribution is obtained for any model in any period, as well as different portfolio optimization results. The view's uniform ranges are centered on the asset class expected returns and have width equal to a standard deviation.

A practitioner can have *K* = *N* subjective views on linear combinations of the market **M**. These linear combinations are represented by a (*K* × *N)* dimensional pick matrix **P**: the generic *k-th* row of the pick matrix determines the weights of the *k-th* view. As I express one absolute view for each asset class at any time over the out-of-sample period, the pick matrix **P** reads:

$$(68) \quad \mathbf{P} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The resulting invertible matrix:

$$(69) \quad \overline{\mathbf{P}} \equiv \begin{pmatrix} \mathbf{P} \\ \mathbf{P}^{\perp} \end{pmatrix}$$

defines the view-adjusted market coordinates. $\mathbf{P}^{\perp}$ can be any *(N - K)* × *N* matrix such that **P** is invertible. However, for computational efficiency and accuracy Meucci suggests to impose that the rows of $\mathbf{P}^{\perp}$ be orthogonal and that their norm equal one. In other words, the *N*-dimensional random vector:

$$(70) \quad \mathbf{V} \equiv \overline{\mathbf{P}}\mathbf{M}$$

is fully equivalent to the market **M**. Then I rotate the market into the views coordinates as in (70). In terms of the empirical panel representation of the market prior (64) this operation corresponds to a simple matrix multiplication:

$$(71) \quad \mathcal{V} \equiv \mathcal{M}\overline{\mathbf{P}}'$$

Each row of the (*J* × *N)* panel $\mathcal{V}$ is an independent joint realization of the market in the views coordinates. The practitioner's views correspond to statements on the first *K* entries of $\mathbf{V}$. Each of these statements, i.e. the generic *k-th* subjective view, is expressed in terms of a cumulative distribution function (cdf):

$$(72) \quad \hat{F}_k(v) \equiv \mathbb{P}_{subj}\{V_k \leq v\} \quad k = 1, \dots, K$$

The prior distribution (63) also implies a distribution for each view, which is represented by its cdf:

(73) $F_k(v) \equiv \mathbb{P}_{prior}\{V_k \leq v\} \quad k = 1, \dots, K$

In general the practitioner's views (72) are different than the respective market implied distributions (73). I thus sort the first *K* columns of the panel (71) and obtain two (*J* × *K)* panels *W* and *C*. The prior cdf (73) of the *k-th* view evaluated at the sorted values is approximated with extreme accuracy by its empirical counterpart:

(74) $F_k(W_{j,k}) \equiv \frac{j}{J+1}$

The generic element $C_{i,k}$ of the panel *C* represents the ranking of the respective entry of the panel $\mathcal{V}$ within its column, normalized by the total number of simulations. Each row of the panel *C* represents a joint simulation of the formula (77). In other words, the ranking panel *C* represents the copula of the views:

(75) $\text{copula } \mathbf{C} \Leftrightarrow \text{ranking } C$

The COP approach resolves this dichotomy by means of opinion pooling techniques. The posterior cdf is defined as a weighted average:

(76) $\tilde{F}_k \equiv c_k \hat{F}_k + (1 - c_k) F_k \quad k = 1, \dots, K$

where the weights c_k represent the confidence in the respective view. Formula (76) defines the posterior marginal distribution of each view. In the COP approach the copula of the posterior distribution of the views is inherited from the prior copula:

(77) $\begin{pmatrix} C_1 \\ \vdots \\ C_K \end{pmatrix} =^{\mathrm{d}} \begin{pmatrix} F_1(V_1) \\ \vdots \\ F_K(V_K) \end{pmatrix}$

The posterior marginal cdf's are approximated with extreme accuracy by their empirical counterparts. From (69) it follows that the (*J* × *K*) panel $\mathcal{F}$ defined as:

(78) $\tilde{F}_{j,k} \equiv c_k \hat{F}_k(W_{j,k}) + (1 - c_k) \frac{j}{J+1}$

represents the posterior marginal cdf’s (76) evaluated in correspondence to the respective entries in the sorted panel *W*. Therefore the posterior joint distribution of the views is defined as follows:

(79) $$\begin{pmatrix} \mathrm{V}_1 \\ \vdots \\ \mathrm{V}_\mathrm{K} \end{pmatrix} \overset{\mathrm{d}}{=} \begin{pmatrix} \tilde{\mathrm{F}}_1^{-1}(\mathrm{C}_1) \\ \vdots \\ \tilde{\mathrm{F}}_\mathrm{K}^{-1}(\mathrm{C}_\mathrm{K}) \end{pmatrix}$$

I can produce a *(J × K)* panel $\tilde{\mathcal{V}}$ of joint scenarios from the posterior distribution of the views (79) by means of the linear interpolator. Finally, to determine the posterior distribution of the market $\mathbf{M} \sim \tilde{f}_\mathrm{M}$ I apply (70) backwards and rotate the views back into the market coordinates:

(80) $$\mathbf{M} =^{\mathrm{d}} \overline{\mathbf{P}}^{-\mathbf{1}}\mathbf{V}$$

where the first *K* entries of $\mathbf{V}$ follow the posterior distribution (79) and the remaining entries are left unaltered. Empirically it is equivalent to rotate these simulations back into the market coordinates as in (80) by means of a simple matrix multiplication:

(81) $$\tilde{\mathcal{M}} \equiv \tilde{\mathcal{V}}\overline{\mathbf{P}}'^{-\mathbf{1}}$$

Each N-dimensional row represents an independent joint scenario from the COP posterior market distribution.

As Meucci (2006) has already summarized, the COP approach starts from a fully generic prior distribution (63) and yields a posterior distribution (81) that reflects the practitioner’s views (72). As in BL, the COP posterior distribution represents a gentle, consistent twist of the original prior distribution, see Meucci (2006): therefore, once fed into an optimization algorithm, it gives rise to sensible allocations. In order to obtain the posterior it is needed to perform the marginalizations (73), the copula factorization (77) and the quantile inversions (79). None of these operations is feasible in closed analytical form, except in trivial cases. Nevertheless, it is straightforward to perform these operations numerically, as I proceed to show.

The COP posterior market distribution obtained for each period during the out-of-sample is then used into a mean-variance and a conditional value-at-risk (CVaR) portfolio optimization.

5.2 Out-of-sample mean-variance asset allocation exercise based on COP and MSVAR(2,1) and VAR(1)

In this subparagraph the results of the mean-variance portfolio optimization based on the COP posterior market distributions are shown. The MSVAR(2,1) and VAR(1) COP posterior multivariate market distributions obtained are then used to compute, for each period, the asset classes expected returns and standard deviations that are fed into a mean-variance portfolio optimization. The same five recursive optimal portfolios as in the in-sample period case have been built over the out-of-sample period.

Table 41 out-of-sample average optimal mean-variance portfolios expected returns based on empirical copula opinion pooling asset classes returns distribution

maximum Sharpe Ratio portfolio with opened lower budget constraint		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0030	0.0024
30%	0.0044	0.0041
50%	0.0156	0.0074
70%	0.0070	0.0071
90%	0.0058	0.0054
maximum Sharpe Ratio portfolio with budget constraint		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0036	0.0038
30%	0.0057	0.0059
50%	0.0338	0.0115
70%	0.0133	0.0082
90%	0.0093	0.0066
optimal weights with capital allocation and risk aversion coefficient = 1		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0049	0.0044
30%	0.0073	0.0070
50%	0.0706	0.0270
70%	0.0261	0.0136

90%	0.0172	0.0107

optimal weights with capital allocation and risk aversion coefficient = 3

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.0063	0.0057
30%	0.0091	0.0082
50%	0.1798	0.0766
70%	0.0631	0.0303
90%	0.0397	0.0210

optimal weights with capital allocation and risk aversion coefficient = 5

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.0068	0.0056
30%	0.0111	0.0093
50%	0.7413	0.3647
70%	0.2509	0.1266
90%	0.1529	0.0790

Table 42 out-of-sample average optimal mean-variance portfolios expected Sharpe ratio based on empirical copula opinion pooling asset classes returns distribution

maximum Sharpe Ratio portfolio with opened lower budget constraint

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.2262	0.1432
30%	0.2262	0.1432
50%	0.1368	0.1209
70%	0.1847	0.1263
90%	0.2084	0.1300

maximum Sharpe Ratio portfolio with budget constraint

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.2911	0.1923
30%	0.2911	0.1923
50%	0.1895	0.1526
70%	0.2226	0.1645
90%	0.2484	0.1710

optimal weights with capital allocation and risk aversion coefficient = 1

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.4008	0.2699
30%	0.4008	0.2699

50%	0.2915	0.2178
70%	0.3128	0.2277
90%	0.3310	0.2372

optimal weights with capital allocation and risk aversion coefficient = 3

view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.5982	0.4052
30%	0.5982	0.4043
50%	0.4734	0.3292
70%	0.4829	0.3354
90%	0.4924	0.3416

optimal weights with capital allocation and risk aversion coefficient = 5

view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	1.1810	0.8065
30%	1.1810	0.8006
50%	0.9804	0.6624
70%	0.9847	0.6666
90%	0.9889	0.6709

A comparison of the out-of-sample expected portfolio returns and the expected Sharpe ratio obtained using the COP posterior market distribution based on both the MSVAR(2,1) and the VAR(1) can be conducted from the observation of Table 41 and Table 42. What emerges from an analysis of Table 41 is that in almost all the cases, given any type of portfolios and any level of confidence in the views, the portfolios based on the MSVAR(2,1) model greatly overperform those based on the VAR(1) model in terms of greater average portfolio expected returns.

If the comparison is conducted in terms of realized Sharpe ratio, as depicted in figure 42, again it can be seen that the portfolios built using the two states model greatly overperform those built using the single state model. It can be seen that the realized Sharpe ratio increases as a function of the risk aversion coefficient and decreases as a function of the confidence level in the views, perhaps as a consequence of the higher variability of the portfolio returns due to the posterior market distribution which greatly depends on the time varying views.

Table 43 out-of-sample average optimal mean-variance portfolios realized returns based on empirical copula opinion pooling asset classes returns distribution

maximum Sharpe Ratio portfolio with opened lower budget constraint		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0067	0.0044
30%	0.0091	0.0078
50%	0.0384	0.0141
70%	0.0148	0.0118
90%	0.0119	0.0072
maximum Sharpe Ratio portfolio with budget constraint		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0068	0.0053
30%	0.0097	0.0098
50%	0.0591	0.0235
70%	0.0228	0.0111
90%	0.0164	0.0077
optimal weights with capital allocation and risk aversion coefficient = 1		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0079	0.0065
30%	0.0102	0.0100
50%	0.0979	0.0250
70%	0.0359	0.0136
90%	0.0238	0.0104
optimal weights with capital allocation and risk aversion coefficient = 3		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0066	0.0069
30%	0.0110	0.0095
50%	0.1606	0.0275
70%	0.0594	0.0167
90%	0.0390	0.0142
optimal weights with capital allocation and risk aversion coefficient = 5		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0073	0.0039
30%	0.0113	0.0096
50%	0.5634	0.1311
70%	0.1931	0.0504
90%	0.1190	0.0339

Table 44 out-of-sample average optimal mean-variance portfolios realized Sharpe ratio based on empirical copula opinion pooling asset classes returns distribution

maximum Sharpe Ratio portfolio with opened lower budget constraint		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.4904	0.3667
30%	0.4871	0.4429
50%	0.3097	0.2878
70%	0.3560	0.2588
90%	0.3925	0.1978
maximum Sharpe Ratio portfolio with budget constraint		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.4779	0.2729
30%	0.3969	0.3395
50%	0.3075	0.3484
70%	0.3546	0.2582
90%	0.4110	0.2138
optimal weights with capital allocation and risk aversion coefficient = 1		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.4181	0.2760
30%	0.3874	0.3295
50%	0.2943	0.1789
70%	0.3229	0.2379
90%	0.3531	0.2443
optimal weights with capital allocation and risk aversion coefficient = 3		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.3453	0.3076
30%	0.3831	0.3083
50%	0.2407	0.0961
70%	0.2675	0.1656
90%	0.2922	0.2156
optimal weights with capital allocation and risk aversion coefficient = 5		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.4058	0.1903
30%	0.3794	0.3089
50%	0.2433	0.1324

70%	0.2503	0.1521
90%	0.2573	0.1693

Overall it can be deduced from Table 43 and Table 44 that the conclusions drawn from the observation of the expected portfolio returns and expected Sharpe ratio are here substantially reproduced even though the benefits derived from the two states model are decreased to some extent.

Table 45 out-of –sample optimal mean-variance portfolio cumulative realized returns based on empirical copula opinion pooling asset classes returns distribution (initial wealth = 1)

maximum Sharpe Ratio portfolio with opened lower budget constraint		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	1.4873	1.2940
30%	1.4893	1.3613
50%	1.5845	1.4492
70%	1.4653	1.4911
90%	1.5343	1.2459
maximum Sharpe Ratio portfolio with budget constraint		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	1.7063	1.5822
30%	1.7534	1.7495
50%	1.8033	1.7699
70%	1.8772	1.7189
90%	1.9139	1.7265
optimal weights with capital allocation and risk aversion coefficient = 1		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	6.0868	2.1617
30%	9.9021	3.5597
50%	-13.0966	2.4393
70%	24.1332	0.1452
90%	-1.52E+09	125.5950
optimal weights with capital allocation and risk aversion coefficient = 3		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	2.2983	1.9021
30%	3.4282	1.8403
50%	5.7637	2.0457

70%	5.5426	2.0094
90%	-144.8532	0.2081

optimal weights with capital allocation and risk aversion coefficient = 5		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	1.9821	1.4811
30%	2.5359	1.5258
50%	3.5870	1.7607
70%	5.8193	2.0566
90%	-66.8947	2.2802

5.3 Out-of-sample conditional value-at-risk (CVaR) asset allocation exercise based on COP and MSVAR(2,1) and VAR(1)

In this subparagraph the results of the conditional value-at-risk portfolio optimization based on the COP posterior market distributions derived from both the MSVAR(2,1) model and the VAR(1) model are shown. In this exercise, for each out-of-sample period the simulated multivariate scenarios are fed into a CVaR portfolio optimization to obtain five different portfolios equally distributed along the time dependent CVaR - expected return efficient frontier; with I being the least risky and V the most risky portfolios. The exercise has been repeated for values of the confidence in the views that span from 0.1 to 0.9. The CVaR probability level has been set to 0.95 for each portfolios.

Table 46 out-of-sample average optimal CVaR portfolios expected returns based on empirical copula opinion pooling asset classes returns distribution

Risk level I		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0000	0.0000
30%	0.0000	0.0000
50%	0.0000	0.0000
70%	0.0000	0.0000
90%	0.0000	0.0000
Risk level II		

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.0016	0.0018
30%	0.0020	0.0019
50%	0.0025	0.0021
70%	0.0031	0.0024
90%	0.0036	0.0027

Risk level III

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.0032	0.0037
30%	0.0039	0.0038
50%	0.0050	0.0043
70%	0.0061	0.0048
90%	0.0073	0.0054

Risk level IV

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.0048	0.0055
30%	0.0059	0.0057
50%	0.0075	0.0064
70%	0.0092	0.0072
90%	0.0109	0.0081

Risk level V

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.0063	0.0073
30%	0.0078	0.0076
50%	0.0100	0.0085
70%	0.0123	0.0096
90%	0.0146	0.0108

Table 47 out-of-sample average optimal CVaR portfolios expected Sharpe ratio based on empirical copula opinion pooling asset classes returns distribution

Risk level I

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.0068	0.0103
30%	0.0146	0.0068
50%	0.0059	0.0013
70%	0.0285	-0.0080
90%	0.0863	0.0476

Risk level II		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.1105	0.0726
30%	0.1412	0.0933
50%	0.1928	0.1245
70%	0.2874	0.1773
90%	0.7093	0.4133

Risk level III		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.1105	0.0725
30%	0.1412	0.0932
50%	0.1928	0.1241
70%	0.2873	0.1766
90%	0.7086	0.4132

Risk level IV		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.1084	0.0665
30%	0.1385	0.0872
50%	0.1869	0.1182
70%	0.2746	0.1685
90%	0.6789	0.4034

Risk level V		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0811	0.0592
30%	0.0978	0.0775
50%	0.1475	0.1060
70%	0.2236	0.1522
90%	0.5637	0.3553

Similarly to what have been shown in the last paragraph, a comparison of the out-of-sample expected portfolio returns and the expected Sharpe ratio obtained using the COP posterior market distribution based on both the MSVAR(2,1) and the VAR(1) can be conducted from the observation of Table 46 and Table 47. What emerges from an analysis of Table 46 is that in almost all the cases, given any portfolio risk level and any level of confidence in the views, the portfolios based on the MSVAR(2,1) model greatly overperform those based on the VAR(1) model in terms of greater average portfolio expected returns.

If the comparison is conducted in terms of realized Sharpe ratio, as depicted in figure 49, again it can be seen that the portfolios built using the two states model greatly overperform those built using the single state model.

Table 48 out-of-sample average optimal CVaR portfolios realized returns based on empirical copula opinion pooling asset classes returns distribution

Risk level I		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0000	0.0000
30%	0.0000	0.0000
50%	0.0000	0.0000
70%	0.0000	0.0000
90%	0.0000	0.0000
Risk level II		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0034	0.0038
30%	0.0033	0.0033
50%	0.0035	0.0030
70%	0.0036	0.0024
90%	0.0035	0.0025
Risk level III		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0068	0.0076
30%	0.0067	0.0065
50%	0.0070	0.0059
70%	0.0071	0.0048
90%	0.0069	0.0049
Risk level IV		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.0100	0.0110
30%	0.0097	0.0098
50%	0.0102	0.0095
70%	0.0103	0.0074
90%	0.0097	0.0071
Risk level V		

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	0.0114	0.0129
30%	0.0111	0.0111
50%	0.0124	0.0124
70%	0.0121	0.0127
90%	0.0119	0.0112

Table 49 out-of-sample average optimal CVaR portfolios realized Sharpe ratio based on empirical copula opinion pooling asset classes returns distribution

Risk level I		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.1255	0.1372
30%	0.2258	-0.1101
50%	0.1185	0.0852
70%	0.2216	0.1135
90%	0.0453	0.2176
Risk level II		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.5084	0.5011
30%	0.4098	0.3938
50%	0.3705	0.3672
70%	0.3615	0.3115
90%	0.3669	0.3070
Risk level III		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.5084	0.2464
30%	0.4098	0.2287
50%	0.3705	0.2163
70%	0.3587	0.1815
90%	0.3590	0.1844
Risk level IV		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.4630	0.3565
30%	0.3825	0.3465
50%	0.3562	0.3441
70%	0.3433	0.2781

90%	0.3231	0.2684
Risk level V		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	0.3080	0.2342
30%	0.2693	0.2403
50%	0.3084	0.2945
70%	0.3027	0.3190
90%	0.2959	0.2846

Overall it can be deduced from Table 49 and Table 50 that the conclusions drawn from the analysis of the expected portfolio returns and expected Sharpe ratio are here substantially confirmed even though the benefits derived from the adoption of the two states model are decreased to some extent, especially for the riskiest portfolio.

Table 50 out-of –sample optimal CVaR portfolio cumulative realized returns based on empirical copula opinion pooling asset classes returns distribution (initial wealth = 1)

Risk level I		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	1.0000	1.0000
30%	1.0000	1.0000
50%	1.0000	1.0000
70%	1.0000	1.0000
90%	1.0000	1.0000
Risk level II		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	1.2237	1.2568
30%	1.2181	1.2127
50%	1.2292	1.1947
70%	1.2353	1.1547
90%	1.2286	1.1583
Risk level III		
view confidence	***MSVAR(2,1)***	***VAR(1)***
10%	1.4925	1.5625
30%	1.4770	1.4621

50%	1.5021	1.4158
70%	1.5111	1.3235
90%	1.4907	1.3306

Risk level IV

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	1.7901	1.8725
30%	1.7484	1.7570
50%	1.7954	1.7212
70%	1.7967	1.5218
90%	1.7370	1.5009

Risk level V

view confidence	*MSVAR(2,1)*	*VAR(1)*
10%	1.8974	1.9781
30%	1.8427	1.8204
50%	1.9953	1.9867
70%	1.9665	2.0340
90%	1.9364	1.8605

5.4 COP asset allocation exercise conclusions

In conclusion it can be said that the adoption of a regime switching model, rather than a single regime model, in the attempt to model the market returns and to provide out-of-sample market forecasts, which the construction of portfolios was based on, in the copula-opinion pooling approach proved to be an overall profitable choice at any level of confidence in the views and with any portfolio optimization method implemented.

CHAPTER 6 – CONCLUSIONS

From the point of view of this dissertation, it is important to note that all the time series displayed significant deviation from the normal distribution benchmark, as evidenced by the statistically significant normality tests results, and then they can be considered not normally distributed. This is a clear indication that the returns on the three asset classes cannot be captured by linear models to reinforce the idea of a regime switching model, which is more flexible in accommodating the mixing

of several empirical distributions. The regime switching model is also supported by the results of normality tests, as all null hypotheses are strongly rejected. All the evidences presented so far suggest that the regime switching estimate of my model appears to be related to the underlying economic fundamentals and business cycle to some extent. From the analysis of the parameter estimates, during regime 1 it seems reasonable to invest in equity, whose average realized return outperforms that of bonds, while in regime 2 holding bonds seems to be more profitable. The presence of a size effect brings to the further conclusion that, regarding the stock market, a strategy that invests in small stocks in regime 1 and large stocks in regime 2 seems to be appropriate and profitable.

From the analysis of the first part of the asset allocation exercise it can be said that while the construction of portfolios based on the two regimes MSVAR(2,1) model estimates does not lead to results that clearly overperform those based on the single regime VAR(1) model estimates in the in-sample framework, it significantly does it in the out-of-sample framework. Therefore the adoption of a sophisticated multi regimes model greatly improves the portfolios performance in the out-of-sample exercise for all types of portfolios constructed.

If the views based on the asset pricing models are taken into account, as in the second part of the asset allocation exercise, it can be said that the adoption of a regime switching model, rather than a single regime model, in the attempt to model the market returns and to provide out-of-sample market forecasts, which the construction of portfolios was based on, in the copula-opinion pooling approach proved to be a profitable choice at any level of confidence in the views and almost at any level of portfolio risk.

The findings reveal that when regime switching is taken into account, the optimal portfolio weights deviate substantially from those that obtain under the single-state under various portfolio selection frameworks. Thus the investment policies are very steep with respect to the state variables, which results in volatile portfolio

policies. As a consequence, portfolio managers should be able to exploit these regime changes to increase profitability.

This paper has illustrated the potential benefit of a portfolio asset allocation exercise based on market returns forecasts produced by a markov regime switching model, which implies its expected returns, their volatility, and their correlations to be time varying and state dependent. On the one hand, I exaggerate the performance of the model due to the fact that I do not take transaction costs into account. However, I underestimate the potential of regime switching models since I did not to estimate the best possible model, indeed I did an extensive model search amongst only vector autoregressive regime-switching models.

To sum up, the regime switching model has proven to be superior than the single regime model due to its performance evaluated by Sharpe ratio and both average expected and average realized portfolio returns. This evidences witness that incorporating the regime switching features into an asset pricing model, which is then used to feed the portfolio optimization process, is an effective and profitable strategy under both favorable and adverse market conditions.

REFERENCES

Ang, A. and Bekaert, G. (2004), “How regimes affect asset allocation”, Financial Analysts Journal, 60, 86–98.

A¨it-Sahalia, Y. and Brandt, M. (2001), “Variable Selection for Portfolio Choice”, Journal of Finance, 56, 1297-1351.

James, J. and Chib, S. (1993), “Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts,” Journal of Business and Economic Statistics 11, 1-15.

Ammann, M. and Verhofen, M. (2006), “The effect of market regimes on style allocation, Financial Markets and Portfolio Management”, 20 (3), 309-337.

Ang, A., and Bekaert, G. (2001), “Regime Switches in Interest Rates,” Journal of Business and Economic Statistics, forthcoming.

Ang, A., and J. Chen, (2001), “Asymmetric Correlations in Equity Portfolios,” forthcoming Journal of Financial Economics.

Ang, A., and Bekaert, G. (2007). Stock return predictability: Is it there? Review of Financial Studies 20 (3), 651–707.

Ang, A.; Bekaert, G. and Liu, j. (2001), “Why Stocks May Disappoint,” working paper, Columbia University.

Ang, A., Timmermann, A., (2011). “Regime changes and financial markets” National Bureau of Economic Research Working Paper Series No. 17182.

Ang, A., and Bekaert, G. (2002), “International Asset Allocation with Regime Shifts”, Review of Financial Studies 15, 1137-1187.

Aslanidis, Nektarios and Christiansen, Charlotte, (2010), “Smooth Transition Patterns in the Realized Stock Bond Correlation”.

Campbell, J. Y., Shiller, R. J. (1988), “Stock prices, earnings, and expected dividends”. The Journal of Finance 43 (3), 661–676.

Campbell, J., and L., Viceira (1999), “Consumption and Portfolio Decisions when Expected Returns are Time Varying”, Quarterly Journal of Economics, 114, 433-495.

Campbell, J., and R., Shiller, (1988), “The Dividend-Price Ratio, Expectations of Future Dividends and Discount Factors”, Review of Financial Studies, 1, 195-227.

Campbell, J., Y., Chan, and L., Viceira, (2003), “A Multivariate Model of Strategic Asset Allocation”, Journal of Financial Economics, 67, 41-80.

Clarke, R. G., and H. de Silva, (1998), “State-Dependent Asset Allocation,” Journal of Portfolio Management,24, 2, 57-64.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977), “Maximum likelihood from incomplete data via the EM algorithm”, Journal of the Royal Statistical Society. Series B. Methodological 39(1), 1–38. With discussion.

Edwards, S. and Susmel, R. (2001), "Volatility Dependence And Contagion In Emerging Equity Markets," Journal of Development Economics, v66(2,Dec), 505-532.

Giot, P. and Petitjean, M, “Short-Term Market Timing Using the Bond-Equity Yield Ratio” (October 1, 2006). European Journal of Finance, 15(4), (2009).

Goetzmann, W., and P., Jorion, (1993), “Testing the Predictive Power of Dividend Yields”, Journal of Finance, 48, 663-679.

Guidolin, M. and Ria, F. (2010), “Regime shifts in Mean-variance efficient frontiers: some international evidence”. Working paper 2010-040B Federal Reserve Bank of St Louis, pp 1–47.

Guidolin, M. and Timmermann, A. (2009), “Forecasts of US Short-Term Interest Rates: A Flexible Forecast Combination Approach”, Journal of Econometrics, 150, 297-311.

Guidolin, M. and Timmermann, A. (2006a), “An Econometric Model of Nonlinear Dynamics in the Joint Distribution of Stock and Bond Returns”, Journal of Applied Econometrics, 21, 1-22.

Guidolin, M. and Timmermann, A. (2006b), “Term Structure of Risk under Alternative Econometric Specifications”, Journal of Econometrics, 131, 285-308.

Guidolin, M. and Timmermann, A. (2005), “Economic implications of bull and bear regimes in UK stock and bond returns” The Economic Journal 115 (500), 111–143.

Guidolin, M. and Timmermann, A. (2007), “Asset allocation under multivariate regime switching”, Journal of Economic Dynamics and Control 31 (11), 3503–3544.

Guidolin, M. and Timmermann, A. (2008a), “International asset allocation under regime switching, skew, and kurtosis preferences”,Review of Financial Studies 21 (2), 889–935.

Guidolin, M. and Timmermann, A. (2008b), “Size and value anomalies under regime shifts”, Journal of Financial Econometrics 6 (1), 1–48.

H.M. Markowitz, (1952), Portfolio selection, The Journal of Finance., 7, 77–91.

Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," Econometrica, 57, 357-384.

Hamilton, J.D. (1996), "Specification Testing in Markov-Switching Time Series Models," Journal of Econometrics, 70, 127-157.

Hamilton, J.D. and G. Lin (1996), "Stock Market Volatility and the Business Cycle," Journal of Applied Econometrics, 11, forthcoming.

Tu, J. (2010), “Is regime switching in stock return important in portfolio decision?” Management Science, (56), 1198–1215.

Kim, Chang Jin (1994), “Dynamic Linear Models with Markov-Switching,” Journal of Econometrics 60, 1-22.

Kim, Chang Jin, and Charles R. Nelson (1999), “State-Space Models with Regime Switching”, Cambridge, Massachusetts: MIT Press.

Krolzig, Hans-Martin (1997), “Markov-Switching Vector Autoregressions : Modelling, Statistical Inference, and Application to Business Cycle Analysis”, Berlin: Springer.

Zhao, Z. (2010), “Portfolio Selection with Jumps under Regime Switching,” International Journal of Stochastic Analysis, vol. 2010, Article ID 697257, 22 pages.

Longin, F., and Solnik, B. (2001), “Correlation Structure of International Equity Markets During Extremely Volatile Periods,” Journal of Finance, 56, 2, 649-676.

Mergner, Sascha ; Bulla, Jan ; Sesboue, Andre ; Chesneau, Christophe (2010), “Markov-switching Asset Allocation: Do Profitable Strategies Exist?”.

Meucci, A. (2005), “Beyond Black-Litterman: Views on Non-Normal Markets”, New York: Lehman Brothers, Inc.

Meucci, A. (2005), “Risk and Asset Allocation”, New York: Lehman Brothers, Inc.

Meucci, A. (2006a), ”Beyond Black-Litterman in Practice: A Five-Step Recipe to Input Views on Non-Normal Markets“, Risk Magazine, September 2006, pp. 114-119.

Mulvey J.M., Zhao Y. (2010), .“An Investment Model via Regime-Switching Economic Indicators”.

Perez-Quiros, G. and Timmermann, A. (2000), “Firm Size and Cyclical Variations in Stock Returns”, Journal of Finance, 55, 1229-1262.

Pesaran, M. H., Timmermann, A., (1995), “Predictability of stock returns: Robustness and economic significance”. The Journal of Finance 50 (4), 1201–1228.

Pettenuzzo, D., Timmermann, A., (2011), “Predictability of stock returns and asset allocation under structural breaks”, Journal of Econometrics 164 (1), 60–78.

Ramchand, L. and Susmel, R. (1997) “Volatility and Cross Correlation Across Major Stock Markets”.

Rockafellar, R., Uryasev, S. (2000), “Optimization of Conditional Value-at-Risk”, USA: University of Florida, University of Washington.

Timmermann, A. (2000), “Moments of Markov Switching Models,” Journal of Econometrics 96, 75-111.

Vo, Huy Thanh and Maurer, R, (2013), “Dynamic Asset Allocation under Regime Switching, Predictability and Parameter Uncertainty”.